



# Research patterns in communication (2009–2019): testing female representation and productivity differences, within the most cited authors and the field

Manuel Goyanes<sup>1,2</sup> · Márton Demeter<sup>3</sup> · Aurea Grané<sup>4</sup> · Tamás Tóth<sup>3</sup> · Homero Gil de Zúñiga<sup>2,5,6</sup>

Received: 30 July 2021 / Accepted: 21 October 2022 / Published online: 15 November 2022  
© The Author(s) 2022

## Abstract

This study compares the share of male/female as first authors, the growth of authors per paper, and the differences in publication productivity in the last decade of the most cited authors versus the field of communication (i.e., a representative sample of papers published in the field of communication). Results indicate that there are significantly more female first authors in the field than a decade ago, but their proportion among the most cited authors has not grown at a similar pace. Likewise, the number of authors per paper has significantly increased in the field, but not among the most cited authors, who, in turn, publish significantly more papers than the field, both in 2009 and 2019. And not only that, the productivity gap between the most cited authors and the field has substantially increased *between* the span of this decade. Theoretical implications of these findings and suggestions for future studies are also discussed.

**Keywords** Communication science · Communication research · Research patterns · Productivity

---

✉ Manuel Goyanes  
mgoyanes@hum.uc3m.es

<sup>1</sup> Department of Communication, Universidad Carlos III de Madrid, Getafe, Spain

<sup>2</sup> Democracy Research Unit, Department of Political Sciences, University of Salamanca, Salamanca, Spain

<sup>3</sup> National University of Public Service, Budapest, Hungary

<sup>4</sup> Department of Statistics, Carlos III University, Getafe, Spain

<sup>5</sup> Media Effects Research Lab, Film Production & Media Studies Department, Pennsylvania State University, State College, PA, USA

<sup>6</sup> Facultad de Comunicación y Letras, Universidad Diego Portales, Santiago, Chile

## Introduction

The analysis of scholarly collaboration, publication and citation patterns is a global enterprise that entails the quantitative measurement of a myriad of scientific phenomena (Heesen & Romejin, 2019). However, like any other social subsystem, science also might be influenced by structural and individual factors, such as the privilege of some languages over the others (Curry & Lillis, 2018; Goyanes, 2020), or differences in terms of research competences and resources. The analysis of such potential differences is paramount to better understand how global knowledge production unfolds and tend to be common topics within the scientometric tradition (Gonzalez-Brambila et al., 2016; Iyandemye & Thomas, 2019). When it comes to the sociological and bibliometric research in communication studies, prior research has so far focused on a plethora of related issues, such as sex bias in the academia (Knobloch-Westerwick et al., 2013), citation networks (Chan & Torgler, 2020), the uneven representation of sex and geographical region on editorial boards (Goyanes & de-Marcos, 2020; Lauf, 2005), or the evolving nature of research patterns across (sub)fields (Freelon, 2013). This paper complements these research endeavors.

Specifically, the objective of this study is twofold. First, it aims to examine the proportion of male/female first author among the most cited authors and in the field of communication, while also testing the number of authors per paper at two comparable but different points in time: 2009 and 2019. Second, the study aims to systematically compare the publication patterns of the most cited authors versus the field for these two years. Particularly, we investigate whether the most cited scholars in 2019 are more productive than the most cited scholars in 2009, and if so, to what extent this growth significantly differs from that of the field. With the examination of these research patterns, our study seeks to move forward the scientometric analysis of communication scholarship in two meaningful ways. First, by investigating the proportion of male/female first authors in both the group of the most cited authors and the field at two different points in time, we seek to clarify whether being male or female has a potential impact on the likelihood of being cited over time. Although the study does not causally determine the individual or environmental factors affecting the likelihood of being a highly cited author, it offers an insightful picture of the field state and evolution. Second, as prior scholarship has not yet considered the empirical comparison of these two sets of samples (most productive vs the field), there is scattered empirical evidence examining the potentially different productivity behaviors of the most cited authors and the field. For this endeavor, we focus on the field of communication, which has a strong tradition of examining research patterns (Demeter & Goyanes, 2021).

## Sex differences in research

According to several scholars, the career progress of female scholars might be burdened by many interconnected, structural factors (Lee & Ellemers, 2019). Sex bias can be observed on both horizontal and vertical dimensions. While the first dimension suggested that female scholars tend to occupy sectors with lower career chances and worse working conditions, the second documented fewer female scholars in top positions (Majcher, 2002). The prestige of the university also influences sex bias, as female scholars are less represented in elite universities. At the top 50 research universities in the US, female scholars held only 31% of the tenured and tenure-track positions in 2019 (NSF, 2019a).

As documented by several studies, scholars argue that, despite all efforts to reduce sex bias in science, women are still “dropping out all the way long to the top rank positions and their share decreases rapidly at higher positions, even if we take into account the cohort effect by which female scholars in academia have a younger age structure” (Majcher, 2002, p. 7). Moreover, Brink and Benschop (2011) showed that the underrepresentation of female scholars cannot be explained by individual motivations alone, as in the humanities and medical sciences the representation of women was significantly higher among job applicants than among recruited scholars and even from the pool of shortlisted applicants females had lower chances to be appointed than males.

In addition, the “individual merit ideology” (Ellemers & van Laar, 2010; Lee & Ellemers, 2019), which tries to explain the situation of female scholars based on individual level competence and merits, has been severely criticized and contested (Barreto & Ellemers, 2015; Major & Kaiser, 2017; Stroebe et al., 2010). For instance, research have shown that even after controlling for many legitimate factors such as age, work experience, performance records, or area of expertise, the academic efforts and achievements of female scholars are less valued than those of their male peers (Ellemers, 2018). Despite equal performance, female scholars are less likely to receive research funding (van der Lee & Ellemers, 2015), have lower chances of being tenured and promoted (Sarsons, 2017), being offered chair positions (Trevino, Gomez-Mejia, Balkin & Mixon, 2015), and are significantly less paid in every career stage than their male colleagues (Shen, 2013). In short, “despite the wide endorsement of individual merit ideology, women in academia have less return on academic investment and achievement than men do (Lee & Ellemers, 2019, p. 66).

Casad et al. (2020) also argue that the underrepresentation of female scholars in academia, especially in STEM, cannot be explained by differences in individual motivations, as the number of female scholars in faculty positions has not increased although more women than before have earned doctorates in STEM (Ginther & Kahn, 2013). From a variety of social factors that hamper the career of women, research has focused on the distribution of family and household responsibilities, in which women are typically more involved (Stack, 2004), the possible careers gaps that can be associated with childcare (Cameron et al, 2016), or different role stereotypes (Eagly et al., 2020; Knoblock-Westerwick et al., 2013). To combine career purposes and childcare, part-time work seems a reasonable path, but it may have a negative influence on both career progression and motivation (Dubois-Shaik & Fusulier, 2017; Majcher, 2002).

Moreover, stereotypically masculine characteristics are more valued in STEM departments, and female scholars frequently report that they face hiring discrimination, experience higher expectations (El-Alayli et al., 2018), and are more likely to be given administrative, mentoring, and service tasks, consequently burdening their own research (Casad et al., 2020). Less time for research negatively influences research output, which can be detrimental to their “success in earning tenure, obtaining research grants and advancing their careers” (Casad et al., 2020, p. 15). Moreover, due to higher expectancies regarding family and childcare, female scholars have lower chances for long travels and international networking, resulting in lower social capital (Collins & Steffen, 2019).

In one of the largest studies on sex differences in sciences, Huang et al (2020) found a complex picture on possible sex bias. Their study (which analyzed the publication history 1.5 million gender identified authors across 83 countries and 13 disciplines from 1955 to 2010), found that despite the emerging number of female scholars, sex differences even increased in both productivity and impact. The authors found that female and male scholars receive a comparable number of annual citations and publish similar number of papers, but

the observed higher drop-out rate and shorter career length for female scholars maintain the overall sex difference.

Moreover, Huang et al (2020) also observed that the sex gap in production is significantly higher for the most productive authors and for scholars working at elite universities, and, while male scholars received significantly more citations than their female peers, this difference is the highest in the case of the most productive authors. As the authors conclude, “despite recent attempts to level the playing field, men continue to outnumber women 2 to 1 in the scientific workforce and, on average, have more productive careers and accumulate more impact” (Huang et al, 2020, p. 4613).

However, controlling for career length and drop-out rates, most sex differences disappear, thus the overall gender gap in publication and impact might be explained by the higher likelihood of losing female scholars over their careers (Huang et al, 2020; Schröder et al., 2021). A Nordic research also found that there are no sex-based differences in hiring practices (Carlsson et al., 2021). However, the same authors emphasize that, while considering comparable performance there is no sex bias in promotion, the significantly lower number of female professors could be explained by male advantages on different levels (such as in monitoring, review boards, or peer-review assessment). In other words, when female scholars develop as competitive portfolios as their male peers, their odds to be promoted is similar.

Despite these differences, many scholars argued that, especially in the Western world, gender equality policies might reduce these gaps. Indeed, several studies found no significant sex differences in promotion after controlling for the number of publications and academic ranks (Jokinen & Pehkonen, 2017). Other experimental studies found that, considering the same lifestyle, recruitment panels might even support female scholars over their male peers (Williams & Ceci, 2015). Another study in fundamental physics found no significant sex differences in citations and hiring opportunities (Strumia, 2021), while the author still found a significant overrepresentation of male scholars. However, the author claims that the male overrepresentation in the field is not a consequence of sex-based discrimination but a result of population level differences.

Similar results were found in Sweden where female scholars have similar odds to be promoted as full professors as their male peers with even lower publication and impact values (Madison & Fahlman, 2020). According to the findings, female scholars are even favored in recruitments and, as authors suggest, this might be a consequence of gender equality policies. Still, the authors report that less than one third of full professors in Sweden are female. All in all, the explanation of the uneven distribution of research production and impact might include, among other important factors, the higher scientific value acquired by the more successful academics, the differences in talent, the resources of research institutions, and many other individual, institutional, and cultural factors.

### **More papers, more authors**

Within the domain of global knowledge production, productivity has been suggested to significantly increase over time (Plume & Weijen, 2014). However, as both the number of (co) authors and the cognitive capacity of academics are finite, it is widely assumed that this publication prosperity is, at least partially, the consequence of the rapidly growing share of co-authored papers (Zdenek, 2018). Several studies have suggested that increasing the number of authors per publication is the main reason for the increase in the total number of papers, although the empirical evidence is inconsistent.

On the one hand, several studies have suggested that the increase of co-authorships is, at least partially, at the cost of sole studies (Goldstein et al., 2005; Greene, 2007; Zetterström, 2004), which, according to some observers, inextricably engenders knowledge fragmentation (Valsiner, 2006). The growing specialization of social research mirrors the productive strategies of hard science, in which multiple authors working at research-intensive labs maximize their research efforts (Kim, 2001). Accordingly, the impact of the individual researcher is diminishing. Zetterström (2004) goes even further and documents questionable practices, such as inviting influential extra authors to add prestige and, ultimately, increasing the odds of publication. Other scholars, in turn, addressed the pressure to publish and the publish or perish paradigm as a fundamental factor to energize productivity (Goyanes & Rodríguez-Gómez, 2018).

On the other hand, a number of scholars argue that it is the complexity and interdisciplinary nature of research that invites scientists to start new collaborations (Greene, 2007). Such collaborations are thought to intensify the theoretical cross-fertilization of disciplines and, clearly, productivity too. Multiple collaborations and research specializations also promote new ways of seeing things, solving research problems that may be difficult to answer with sole authorships. In addition, funding agencies also encourage interdisciplinary and multiple university research projects through the granting process, potentially increasing the number of authors per article. Ultimately, as a vocal strand of academic discourse maintains, research is a teamwork endeavor. While there are divergent assumptions on potential causes of the growing fractional authorship (Plume & Weijen, 2014), there is, in turn, a strong consensus about two salient assumptions of global knowledge production: productivity is becoming the golden standard of research assessments and both the number of published papers, and the share of co-authored studies have substantially risen over time.

## Hypotheses and research question

As the literature suggests, although not without conflicting empirical results (Haslam et al., 2008; Over, 1990), the share of female first authors has raised over time. In the Western world, it is generally assumed that the ratio of female to male scholars has changed in most domains, with less difference between the ratio of sexes in favor of men within the research institutions (Bolzendahl & Myers, 2004). It is plausible that these changes may also transpire across communication scholars, resulting in more female scholars as first authors at both the level of research production and citation (Knobloch-Westerwick & Glynn, 2013). Accordingly, we presume that *(H1) the field of communication has experienced a significant increase in the proportion of a) female first authors in the field and b) female first authors within the group of most cited authors in 2019, compared to 2009.*

Within the literature on global knowledge production there is extensive empirical evidence pointing at an increasing number of co-authored papers at the cost of single-authored ones (Greene, 2007; Zetterström, 2004). Accordingly, we assume that the number of co-authored papers may increase over time within both the group of the most cited authors and the representative set of all authors in communication. Thus, we hypothesize that *(H2) the field of communication has experienced a significant increase in the number of (a) authors per paper in the field and (b) authors per paper within the group of most cited authors in 2019, compared to 2009.*

While there is extensive research on the linkage between quantity and quality in publication trends (Györfy et al., 2020; Larivière & Costas, 2016), previous analyses examined either a general pool of authors (Kurambayev & Freedman, 2020) or top-performing

authors (Fahy, 2018; Lincoln et al., 2012). Thus far, little empirical research has examined both subsamples. Our study directly compares the productivity of top-cited authors with that of the field. In line with the assumption that the most cited authors publish more papers (Sandström & van den Besselaar, 2016), we hypothesize that *(H3) the group of most cited authors publish more papers than (the general trend of authors in) the field in (a) 2009 and (b) 2019.*

The last hypothesis is based on two corresponding assumptions. First, there is ample evidence suggesting that the overall scientific output, measured by the absolute number of published articles, has increased over time (Chang et al., 2020; Zetterström, 2004). Second, it is also well-known that authors with more *impact*, measured by the number of citations, publish more papers than their average peers (Gyórfy et al., 2020; Lariviere & Costas, 2016) since a reasonable pathway for bolstering citations is productivity (Sandstrom & van den Besselaar, 2016). Accordingly, we assumed that these two combined factors result in an even more pronounced increase in the number of published articles for the most cited authors than for the less cited authors. Thus, we predict that *(H4) there has been a statistically significant increase in the number of published papers by the group of most-cited authors with respect to the field in 2019 compared to 2009.*

Finally, our research question relates to the possible association between co-authorship and productivity (as measured by the total number of individual published papers) in the case of the most cited authors. Specifically, based on Greene's (2007) research that found a significant positive association between the number of published articles and the number of co-authors, we empirically test the association between the number of co-authors and productivity. This empirical testing would provide a better contextualization of the differences in publication productivity of the most cited authors in 2019, in contrast to those of 2009. Accordingly, we pose the following research question, *(RQ1) which are the main productivity behaviors within the group of most cited authors in 2009 and 2019, as it is represented by the number of authors per paper and the total number of different authors?*

## Methods

We used SciVal, a platform that works with Scopus data, to list the most cited authors in communication in 2009 and 2019. While SciVal is a relatively new tool, it is widely used to examine research frontiers and evaluate top scholarly performance (Santos et al., 2020). We selected the 50 most cited authors in each period to implement a parsimonious analysis. In restricting our analysis to the 50 most cited authors we followed the methodology of Shapiro (2000) who also considered the 50 most cited authors in his analysis of the citation patterns of legal scholars. As the citation-based distribution of scholars tends to follow a Pareto distribution (Wang & Barabási, 2021), this set of authors represents the “top of the top” with specific properties of citation records that make it interesting to compare them with the field.

We decided to work with these two periods because we aimed to analyze the widest interval possible. For 2019, SciVal works with citation data between 2017 and 2019. Unfortunately, there was no such an option for 2009. Thus, we imported data on published papers from Scopus (between 2007 and 2009) and then computed the list of the most cited authors in SciVal.

For data gathering, first, we collected the names of the most cited authors (different in each period of study, 2009 and 2019) and every co-author's name whom the most cited

authors worked with in the analyzed years. Second, we accumulated the number of co-authors and computed the number of published papers by the most cited author without any co-author. Third, we gathered the proportion of papers in which the most cited author worked with one or more co-authors separately and cumulated the names of every co-author, including the most-cited author, to estimate how many authors worked altogether in the projects where the most cited author was involved. Finally, we coded manually the sex of the most cited authors by looking at their names in their Curriculum Vitae (CVs).

Another dataset based on the Journal Citation Reports (JCR) was built for the analysis of the field of “Communication” in both 2009 (N of articles considered=315) and 2019 (N of articles considered=356). The JCR ranking was selected because it is arguably the most common classification tool in research assessments. In 2009 a total of 1756 articles were published in JCR journals, while in 2019 this number had increased to 4800. Thus, the sample sizes considered are representative of the respective target population and proportional to the number of published papers per year of each journal at 5% margin of error. Special issues were discarded from the analysis as they might introduce bias. After acquiring the proportional sample sizes, we randomly selected articles with the help of a random number generator to create a representative sample that will be eventually coded. Then, we listed the relevant papers, counted and searched the total numbers and names of authors in every published paper from the sample. According to our calculations, 645 and 823 scholars, including the first authors, worked on the selected papers in 2009 and 2019.

## Data analysis protocol

We ran different statistical techniques to test our hypotheses and answer our research question. For H1a and H1b, we compared probability distributions between sex (male/female) and years (2009/2019), for the most cited authors and the field (i.e., a representative sample of papers in the field of communication), independently. Accordingly, we ran a series of Chi-square homogeneity tests ( $\chi^2$ ) with Yate’s correction (2×2 cross-tables). Similarly, for H2a and H2b, we compared probability distributions of the number of authors per paper (one author, two or three authors, and four or more authors) and years (2009/2019). In order to test if there are statically significant differences, we ran Chi-square homogeneity tests ( $\chi^2$ ) and Kendall’s rank correlation coefficients Tau-b ( $\tau_b$ ) and Tau-c ( $\tau_c$ ).

For testing H3a and H3b, we first computed the productivity of the most cited authors (i.e., the number of published papers) in both 2009 and 2019. Secondly, we estimated the papers per author in the field. Conceivably, straightforward estimations can be obtained by computing sample statistics in both 2009 and 2019. However, by estimating the probability distribution of the mean (or the median) by resampling techniques (i.e., bootstrapping), we obtained a more unbiased estimation as resampling does not only provide point estimators, but also confidence intervals. A brief explanation of the methodological protocol is found below.

In this analysis, “PP” is the total number of published papers for a given year and “N” is the number of different authors. The author productivity can be estimated by computing the number of papers within those PP in which they have collaborated, yielding a set of N values, and registering the productivity per author. By bootstrapping the set of N productivities, we get B sets n productivities. In our case, we took B=1,000. Finally, we computed the statistical measure of interest (in our case, the mean) in each of the B groups. This yielded to a sample of B values from which a (1- $\alpha$ )100%-confidence interval can be

composed by considering the percentiles  $\alpha/2$  and  $1-\alpha/2$  of the probability distribution, for  $0 \leq \alpha \leq 1$ .

Finally, for testing H4 we estimated the difference between the average productivity in the field in 2009 and 2019 using resampling techniques. Our methodological protocol was performed as follows. Let us consider  $y_1$  and  $y_2$  the two different years of examination (i.e., 2009 and 2019). We propose to perform a test with a null hypothesis, where represents the mean productivity in the target population in the corresponding year  $y_i$  ( $i=1,2$ ), by using the following test statistic:

$$productivity\_increase = \bar{x}_{y_1} - \bar{x}_{y_2}$$

where  $\bar{x}_{y_i}$  is the sample mean productivity in year  $y_i$  ( $i=1,2$ ). The distribution of the statistic under the null was computed from  $B=1000$  bootstrap samples of sizes  $n_1$  and  $n_2$ , where  $n_i$  is the number of published papers in year  $y_i$ ,  $i=1,2$ . A  $(1-\alpha)100\%$ -confidence interval is obtained by considering the percentiles  $\alpha/2$  and  $1-\alpha/2$  of the probability distribution, for  $0 \leq \alpha \leq 1$ .

Finally, for answering RQ1, we employed a different mechanism for visualizing and testing the potential actions of the most cited authors to increase their productivity levels in 2009 and 2019. First, we plotted their (1) productivity and the total number of different co-authors and their (2) productivity and the number of authors per paper in both years. We also ran a Spearman's correlation and described their productive behavior according to these variables. Finally, we plotted the whole sample distribution. Specifically, we used the "functional boxplot" (Sun & Genton, 2011) of the sample curves. A functional boxplot is a graphical tool that allows visualizing a set of curves or functions highlighting the central bulk of curves and outlying curves, equivalently as a boxplot allows to visualize the distribution of univariate data.

## Results

### Hypothesis 1

Hypothesis H1a predicted a significant increase in the proportion of female first authors within the field in 2019 with respect to 2009. In 2009 the proportion of female first authors was 45%, and in 2009 it had increased to 57%. These observed differences are statistically significant at a 5% significance level ( $\chi^2(1)=8.694$ ,  $p=0.0032$ ). These results are reported in Table 1. H1a was supported.

Hypothesis H1b predicted a significant increase in the proportion of female first authors within the group of most cited authors in 2019 with respect to 2009. In 2009 the proportion

**Table 1** Results of the Chi-square test for H1a

|       | First author sex |        | All authors |
|-------|------------------|--------|-------------|
|       | Male             | Female |             |
| 2009  | 167              | 138    | 305         |
| 2019  | 147              | 194    | 341         |
| Total | 315              | 332    | 646         |

Chi-square test (with Yates correction):  $\chi^2(1)=8.694$ ,  $p=0.0032$



**Table 2** Results of the Chi-square test for H1b

|       | First author sex |        | All authors |
|-------|------------------|--------|-------------|
|       | Male             | Female |             |
| 2009  | 33               | 17     | 50          |
| 2019  | 29               | 21     | 50          |
| Total | 62               | 38     | 100         |

Chi-square test (with Yates correction):  $\chi^2(1)=0.382, p=0.537$

**Table 3** Distribution of the number of authors per paper within the field in 2009 and 2019 and summary statistics

| Number of authors per paper | Percentage in 2009 (315 papers) | Percentage in 2019 (356 papers) |
|-----------------------------|---------------------------------|---------------------------------|
| 1                           | 43.3%                           | 31.7%                           |
| 2                           | 30.9%                           | 32.6%                           |
| 3                           | 14.6%                           | 21.3%                           |
| 4                           | 6.1%                            | 7.6%                            |
| 5                           | 2.5%                            | 4%                              |
| 6 or more                   | 2.5%                            | 2%                              |
| Total                       | 100.0%                          | 100.0%                          |
| Mean                        | 2.03                            | 2.31                            |
| Median                      | 2                               | 2                               |
| 40th percentile             | 1                               | 2                               |
| Maximum                     | 8                               | 17                              |
| SD                          | 1.28                            | 1.48                            |

of female first authors was 34%, and in 2019 it was 42%. However, these observed differences are not statistically significant at a 5% significance level ( $\chi^2(1)=0.382, p=0.537$ ). The results are shown in Table 2. H1b was not supported.

### Hypothesis 2

H2a predicted a significant increase in the number of (a) authors per paper within the representative field and (b) authors per paper within the group of most cited authors in 2019 with respect to 2009. These distributions are reported in Table 3.

In order to test if these differences are statistically significant, we conducted a Chi-square test of homogeneity ( $\chi^2(1)=9.674, p=0.008$ ) (see Table 4). Thus, we conclude that these populations do not follow the same probability distribution. Moreover, Kendall’s Tau-b and Tau-c lead us to conclude that the number of authors per paper has increased over time. Therefore, H2a is supported.

For H2b, the most cited authors in 2009 published a total of 165 papers, co-authored by 316 different authors, while the most cited authors in 2019 published a total of 280 papers, co-authored by 499 different authors. These distributions are shown in Table 5. We observe a slight increase in mean from 3.32 authors in 2009 to 3.55 in 2019. The median of authors is 3 in both cases. However, the 64th percentile indicates that in 2009 the number of papers authored by three or fewer authors was greater than those in 2019. Indeed, when

**Table 4** Results of the Chi-square test for H2a

|  | Number of authors per paper |                      |                      |       | All papers      |
|--|-----------------------------|----------------------|----------------------|-------|-----------------|
|  | One author                  | Two or three authors | Four or more authors |       |                 |
| 2009                                     | 136                         | 143                  | 35                   |       | 315             |
| 2019                                     | 113                         | 192                  | 51                   |       | 356             |
| Total                                    | 249                         | 335                  | 86                   |       | 670             |
| Chi-square test and association measures |                             |                      |                      |       |                 |
|  |                             |                      |                      | value | <i>p</i> -value |
| Chi-square statistic                     |                             |                      |                      | 9.674 | 0.008           |
| Kendall's Tau-b                          |                             |                      |                      | 0.111 | 0.003           |
| Kendall's Tau-c                          |                             |                      |                      | 0.120 | 0.003           |

**Table 5** Distribution of the number of authors per paper within the group of most cited authors in 2009 and 2019 and summary statistics

| Number of authors per paper | Percentage in 2009 (165 papers) | Percentage in 2019 (280 papers) |
|-----------------------------|---------------------------------|---------------------------------|
| 1                           | 10.3%                           | 11.8%                           |
| 2                           | 24.2%                           | 21.4%                           |
| 3                           | 30.3%                           | 26.1%                           |
| 4                           | 18.8%                           | 19.3%                           |
| 5                           | 6.7%                            | 8.2%                            |
| 6 or more                   | 9.7%                            | 13.2%                           |
| Total                       | 100.0%                          | 100.0%                          |
| Mean                        | 3.32                            | 3.55                            |
| Median                      | 3                               | 3                               |
| 64th percentile             | 3                               | 4                               |
| Maximum                     | 11                              | 10                              |
| SD                          | 1.85                            | 2.08                            |

the cumulated percentages are computed, the cumulated percentage at 3 is 64.8% in 2009, while this value is 59.3% in 2019. These observed differences are not statistically significant at a 5% significance level ( $\chi^2(1)=1.130, p=0.288$ ). The results are shown in Table 6. H2b was not supported.

**Hypothesis 3**

H3 predicted that the group of most cited authors publish more papers than the general trend of authors in the field in (a) 2009 and (b) 2019. For H3a, an estimation of the average number of papers per author in the field is 1.022 (SD of 0.148) and the median is 1. A bootstrap confidence interval for the field in 2009 is developed by taking  $PP=315, N=624$  and  $B=1000$ . In particular, we consider  $\alpha=0.05$  and get the 95%-confidence interval as (1.0128, 1.0353).

For the most cited authors in 2009, the estimated average number of papers per author was 3.32, which is not included in the previous 95% CI. Accordingly, with a 5% significance level, the productivity of this group of authors does not follow that of the field. Additionally, since their estimated average productivity is greater than the upper bound of the 95%-CI, our results indicate that, on average, the most cited authors published more papers than the general trend of authors in the field in 2009. H3a was supported.

**Table 6** Results of the Chi-square test for H2b

|       | Number of authors per paper |                      | All papers |
|-------|-----------------------------|----------------------|------------|
|       | Three or less authors       | Four or more authors |            |
| 2009  | 107                         | 58                   | 165        |
| 2019  | 166                         | 114                  | 280        |
| Total | 273                         | 172                  | 445        |

Chi-square test (with Yates correction):  $\chi^2(1)=1.130, p=0.288$

For H3b, the average number of papers per author in the field in 2019 is 1.028 (SD of 0.164) and the median is 1. A bootstrap confidence interval for the field in 2019 is developed by taking  $PP=356$ ,  $N=792$  and  $B=1000$ . In this case, we get the 95%-confidence interval as (1.0152, 1.0366). The estimated average number of papers per author is 3.55, which is not included in the previous 95%-CI. Thus, with a 5% significance level, the productivity of the most cited authors do not follow that of the field. Additionally, since their estimated average number of papers per author is greater than the upper bound of the 95%-CI, we conclude that the most cited authors publish more papers than the general trend of authors in the field in 2019. Therefore, H3b was supported.

## Hypothesis 4

H4 predicted a significant increase in the number of published papers by the group of most cited authors with respect to that of the field in 2019 compared to 2009. In order to test this hypothesis, we separately tested the difference across time within each group (the field and the most cited authors) to finally compare them. We started by analyzing the difference between the average productivity in the field in 2019 and then, in 2009. A straightforward estimation is given by the difference of the corresponding averages, which yields 0.005. Additionally, we obtained a bootstrap confidence interval for the difference in average productivity in the field in 2019 and 2009 (at a 95% confidence level) by taking  $n_1=792$  and  $n_2=624$  and  $B=1000$  and obtain  $(-0.0200, 0.0136)$ . Since zero is included in the 95%-CI, we conclude that the increase in the average number of papers per author in the field is not statistically significant at a 5% significance level.

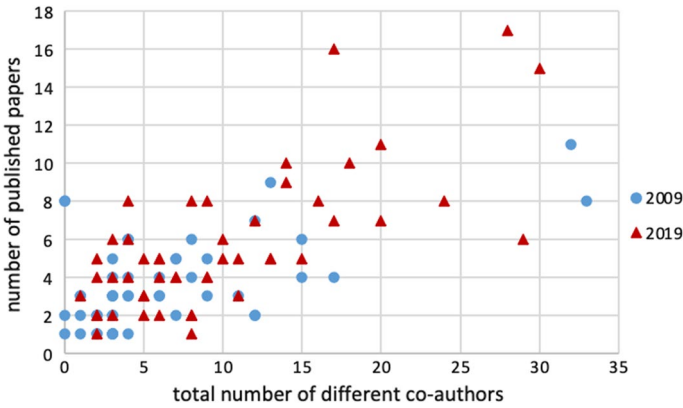
Going back to the most cited authors in 2019 and 2009, the estimated difference for their average productivity is 0.23, which is not included in the previous 95%-CI. Indeed, the estimated difference is greater than the upper bound of the 95%-CI. Thus, our findings indicate that the increase in the average productivity of the most cited authors is statistically significant and greater than that of the field. Therefore, H4 was supported.

## Research question 1

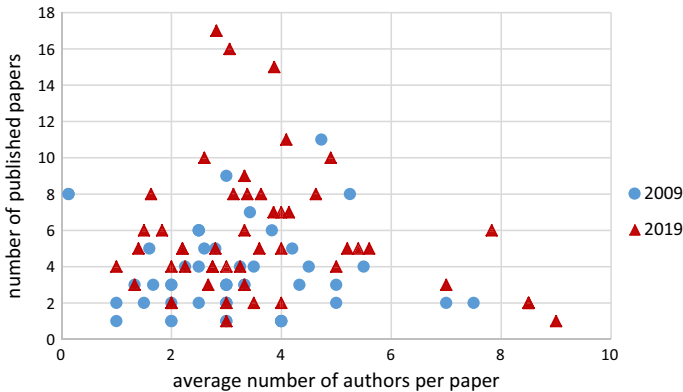
Figure 1 illustrates the number of published papers versus the total number of different co-authors among the most cited authors, where a trend can be observed (RQ1): the higher the number of different co-authors, the higher the number of published papers. Indeed, in 2009, Spearman's correlation coefficient between this pair of variables was 0.652 and in 2019 was 0.778 (both being statistically significant at a 5% significance level).

Looking at the two time periods, a change in productivity behavior can be detected. The median of the total number of co-authors was 4 in 2009, while the median value in 2019 was 8. That is, in 2019, 50% of the most cited authors collaborated with eight or more authors, which is twice the corresponding value in 2009. As a result, the median of published papers in 2019 reached 5, whilst in 2009 this value was 3. Thus, it seems that doubling the number of co-authors is a successful way to increase productivity.

The average number of authors per paper among the most cited authors ranged from 1 to 7.50 in 2009, with a median of 3. In 2019, the average number of authors per paper among the most cited authors took greater values, ranging from 1 to 9, with a median of 3.3. Figure 2 shows the number of published papers versus the average number of authors per paper among the most cited authors, where three productivity behaviors can be sketched.



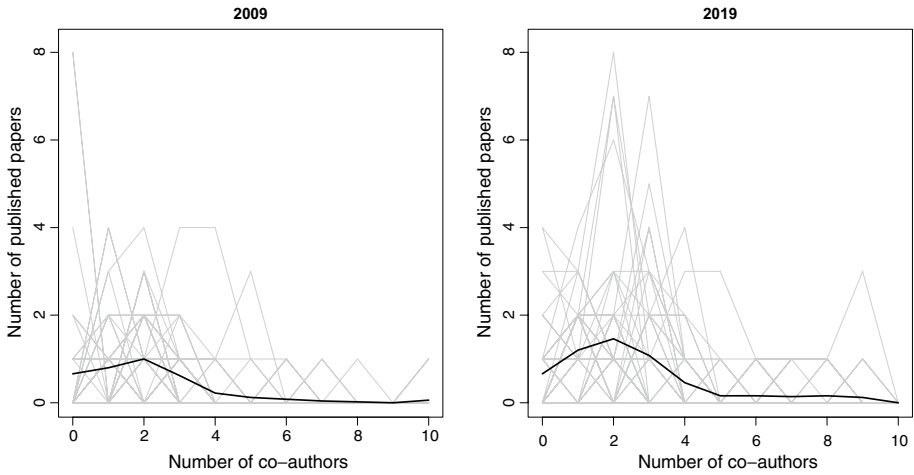
**Fig. 1** Number of published papers versus total number of different co-authors among the most cited scholars



**Fig. 2** Number of published papers versus average number of authors per paper among the most cited scholars

The first group of three authors with the highest level of productivity (more than 14 papers) and an average number of co-authors between 2 and 4. The second group of six authors with low levels of productivity (in general, less than 4 papers) and the highest average of co-authors per paper (larger than 7). The rest of the authors have productivity levels ranging from 1 to 11 and an average number of co-authors from 1 to 5 and can be classified into several groups. For example, we may consider that those authors who publish more than 6 papers per year have a high productivity level, while those with 4–6 papers and those with less than 2 papers per year can have a medium and low productivity levels.

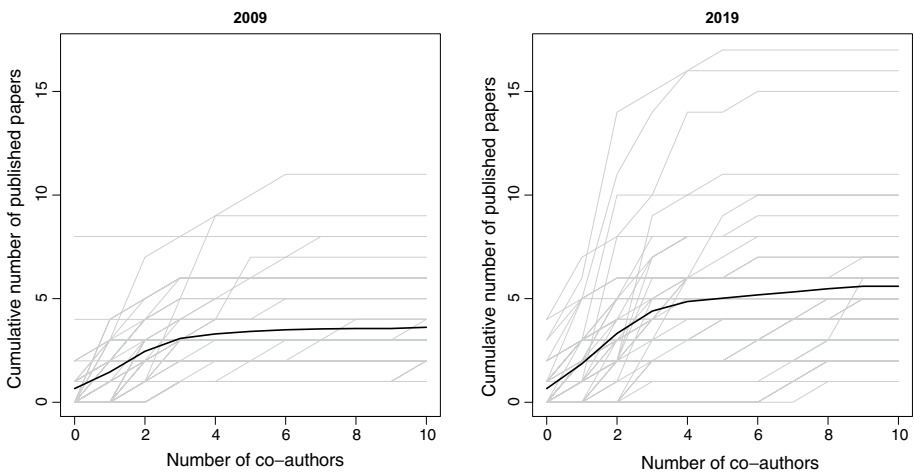
An analysis of the number of publications for each author as a function of the number of collaborators also yields relevant conclusions. Figure 3 displays data as curves where each line represents the publication activity of one author in each of the analyzed years. The black line stands for the pointwise mean: at each given number of co-authors, the corresponding point in this mean line is the average number of publications with that given number of co-authors over every author in the sample. Comparing the sample



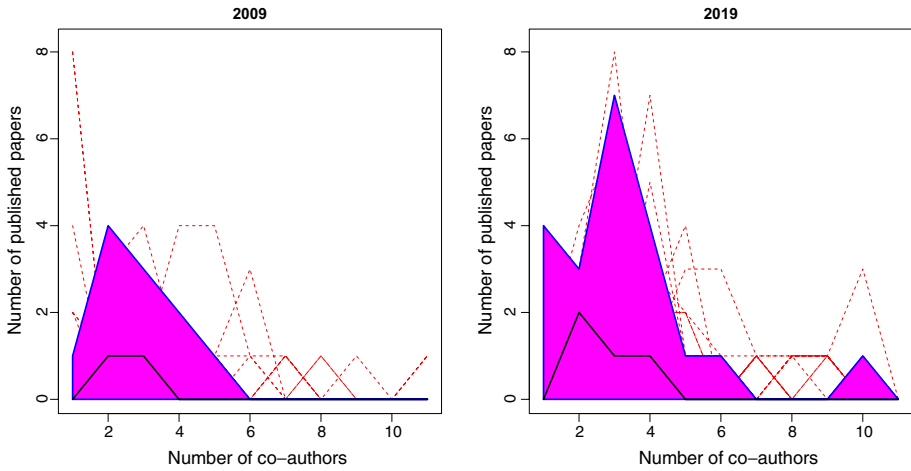
**Fig. 3** Number of publications as a function over the number of co-authors among the most cited scholars. Each line represents an author. The black line stands for the mean pattern

graphics for 2009 and 2019, we observe some differences, with slightly more publications and more collaborators in 2019 than in 2009.

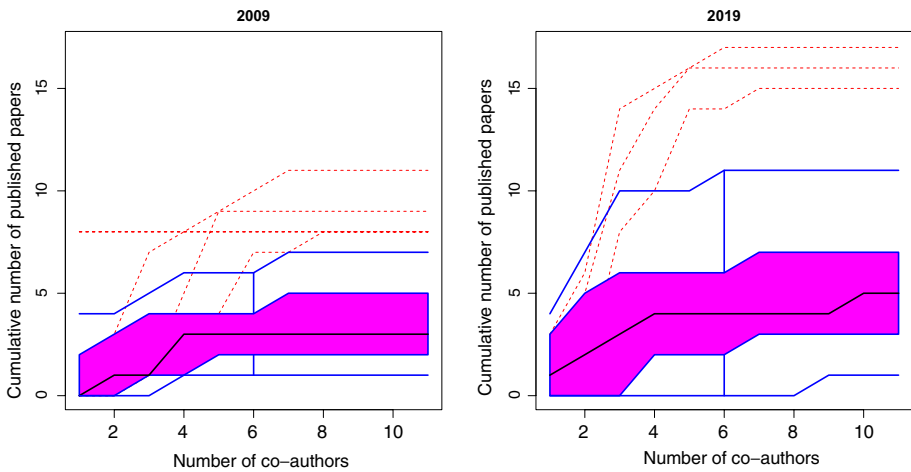
However, the two mean curves are visually different from 0 to 4 co-authors, but then very similar for 4 or more co-authors. If we look at the bigger picture, not focusing on the number of publications for each number of co-authors, but the number of publications up to a given number of co-authors, we look at the cumulative number of papers for each author as a number of its co-authors. The corresponding sample curves and means are presented in Fig. 4.



**Fig. 4** Cumulative number of publications as a function over the number of co-authors among the most cited scholars. Each line represents an author. The black line stands for the mean pattern



**Fig. 5** Functional boxplots of the number of published papers as functions of the number of collaborators among the most cited scholars



**Fig. 6** Functional boxplots of the number of cumulative published papers as functions of the number of collaborators among the most cited scholars

Another interesting analysis is to focus not only on the mean behavior, but also the whole sample distribution. Figures 5 and 6 represent the functional boxplots for the samples of curves in Figs. 3 and 4, respectively.

The pink area in these plots represents the region in which 50% of the most “central” curves lie, whereas the red lines stand for atypical curves. Blue lines are the equivalent of the “whiskers” in a standard boxplot. Finally, the black line stands for the “functional median” or most central curve, which represents the central pattern of the curves in the sample. As opposed to the pointwise mean, the median curve is one of the curves of the sample, so it represents a real trajectory, and it is not smoothed.

In Fig. 5, we appreciate again the growing number of papers and co-authors in 2019 with respect to 2009. We can also see that a significant number of authors have different behavior than the central one (red lines represent outlying trajectories) in both years. When we look at the functional boxplot of the cumulative number of published papers, now the behavior is more homogeneous across authors, and there are only three or four atypical authors (depending on the year) that outstand the rest with significantly higher production.

## Discussion and conclusions

This study provides several inter-related contributions to move forward the growing scientometric literature on communication. First, at the field level, our findings indicate that the share of female first authors (57%) is greater in 2019 than the share of their male counterparts (43%). However, among the top-cited authors in 2019, male first authors still outperform female authors with a share of 58%. Additionally, we found that the female's share was 45 percent in the field in 2009, while their share was 34 percent among the most cited authors. In 2019, the overall female contribution raised to 57 percent, and their share among the top authors was 42 percent. This suggests that, while the average share of female authors in the field has significantly increased, their proportion among the most cited ones remained the same ( $34/45=0.73$ ;  $43/57=0.74$ ). This outcome at different levels suggests a slow progression within female authors' ratio among top-cited authors that are more visible at the field frontier than among the most cited authors.

Second, our study revealed that the average author/paper quotient have raised over time in the field, but not among the most cited authors. Notwithstanding, we found a significant correlation between the number of co-authors and the number of published papers. These two findings suggest that the most cited authors are those who can maximize the number of published papers by working with more authors, but without raising the average number of authors per paper. This trend indicates that a growth in the number of co-authors boosts productivity only to a certain point, after which productivity gets stagnant or even decreases if the number of co-authors per paper increases. Nonetheless, it is important to mention that other factors, not considered in this study, may also affect productivity. For instance, scholars with more funding might be approached by more scholars who aim to participate in the funded research in exchange for authorship and wage. Additionally, more motivated researchers might produce more papers and they might also attract a larger number of colleagues who strive to join such research ambitions (Zdenek, 2018).

Our study also provides illustrative insight to the scholarly discussion over the relationship between quantity and quality (Chang et al., 2020; Zetterström, 2004). First, we found that top-cited authors publish more papers than the field in both 2009 and 2019. Second, we found that top-cited authors have risen their annual productivity in the last 10 years relative to that of the field. All together, these findings clearly indicate that the productivity gap between top cited authors and the field is increasing over time. Similarly, our findings are consistent with the publish or perish paradigm significantly increasing the number of published papers (Goyanes & Rodríguez-Gómez, 2018).

We found notable differences in publication productivity between the field and the most cited authors. In short, the number of papers per scholar have increased for both groups (the field and the most cited authors), but the number of authors per paper have only increased in the field. Accordingly, adding the number of co-authors seems to be a good way to boost productivity. Still, in the case of the most cited authors, productivity, as



related to the number of co-authors per paper seem to follow a Bell curve: the most productive authors are those with papers with 2–4 co-authors and both authors with less than 2 and more than 4 typically underperform them. It follows that raising the number of authors is rewarding only as the authors per paper quotient reach 4, while the behavior of involving more co-authors might reduce productivity. Overall, our empirical examination opens new avenues of research to better understand research productivity and sex differences in global science (Knobloch-Westerwick et al., 2013; Author, XXXXX; Gliboff, 2018).

## Limitations

This study has some limitations that should be addressed by future research. The random sample of papers in the field for both 2009 and 2019 is representative of the total population of papers published, but not representative of the authors that published a paper during these years. Accordingly, the number of papers in the sample for an author may be lower or equal than the total number of papers published by that author. The value of  $N_{authors-year}$  is difficult to obtain since it would imply accessing each and every published article to account for all the authors. Then, the best approximation we can get implies using the empirical average from the sample of articles, even if we acknowledge the potential limitation of this technique.

Finally, as we decided to analyze the top 50 most cited authors, we can assume that the trends would be different if we choose a narrower or a wider selection. Following the Pareto distribution, citations tend to be accumulated by a limited set of authors, while regular scholars have a limited number of publications only. While in selecting the 50 most cited authors, we followed Shapiro's (2000) analysis, we should admit that different sampling technics and different sample sizes might influence the results.

**Acknowledgements** A prior version of this article was sent to 2021 AEJMC Conference.

**Funding** Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. No funding was received for conducting this study.

## Declarations

**Conflict of interest** Author(s) declare no conflict of interest in the development of this study.

**Ethical approval** Author(s) declare following all ethical rules for conducting this study.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Barreto, M., & Ellemers, N. (2015). Detecting and experiencing prejudice: New answers to old questions. *Advances in Experimental Social Psychology*, *52*, 139–219.
- Bolzendahl, C. I., & Myers, D. J. (2004). Feminist attitudes and support for gender equality: Opinion change in women and men, 1974–1998. *Social Forces*, *83*, 759–789.
- Brink, M., & Benschop, Y. (2011). Slaying the seven-headed dragon: The quest for gender change in academia. *Gender, Work & Organization*, *19*(1), 71–92.
- Cameron, E. Z., White, A. M., & Gray, M. E. (2016). Solving the productivity and impact puzzle: Do men outperform women, or are metrics biased? *BioScience*, *66*(3), 245–252.
- Carlsson, M., Finseraas, H., Midtbøen, A. H., & Rafnsdóttir, G. L. (2021). Gender bias in academic recruitment? Evidence from a survey experiment in the Nordic region. *European Sociological Review*, *37*(3), 399–410.
- Casad, B. J., Franks, J. E., Garaski, C. E., Kittleman, M. M., Roesler, A. C., Deidre, Y. H., & Petzel, Z. W. (2020). Gender inequality in academia: Problems and solutions for women faculty in STEM. *Journal of Neuroscience Research*, *99*(1), 13–23.
- Chan, H. F., & Torgler, B. (2020). Gender differences in performance of top-cited scientists by field and country. *Scientometrics*, *125*, 2421–2447.
- Chang, Y., Chen, D., & Huang, M. (2020). Discovering types of research performance of scientists with significant contributions. *Scientometrics*, *124*, 1529–1552.
- Collins, R., & Steffen, N. (2019). Hidden patterns: Using social network analysis to track career trajectories of women STEM faculty. *Equality, Diversity and Inclusion*, *38*(2), 265–282.
- Curry, M. J., & Lillis, T. (2018). *Global academic publishing, Policies perspectives, and pedagogies*. Bristol: Multilingual Matters.
- Demeter, M., & Goyanes, M. (2021). A world-systemic analysis of knowledge production in international communication and media studies: the epistemic hierarchy of research approaches. *The Journal of International Communication*, *27*(1), 38–58.
- Dubois-Shaik, F., & Fusulier, B. (2017). Understanding gender inequality and the role of the work/family interface in contemporary academia: An introduction. *European Educational Research Journal*, *16*(2–3), 99–105.
- Eagly, A. H., Nater, C., Miller, D. I., Kaufmann, M., & Sczesny, S. (2020). Gender stereotypes have changed: A cross-temporal meta-analysis of US public opinion polls from 1946 to 2018. *American Psychologist*, *75*(3), 301–315.
- El-Alayli, A., Hansen-Brown, A. A., & Ceynar, M. (2018). Dancing back-wards in high heels: Female professors experience more work demands and special favor requests, particularly from academically entitled students. *Sex Roles*, *79*(3–4), 136–150.
- Ellemers, N. (2018). Gender stereotypes. *Annual Review of Psychology*, *69*, 275–298.
- Ellemers, N., & Van Laar, C. (2010). Individual mobility: The opportunities and challenges members of devalued groups encounter when trying to avoid group-based discrimination. In J. F. Dovidio, M. Hewstone, P. Glick, & V. Esses (Eds.), *Handbook of prejudice, stereotyping, and discrimination* (pp. 561–576). Sage.
- Fahy, D. (2018). The laureate as celebrity genius: How scientific American's John Horgan profiled nobel prize winners. *Public Understanding of Science*, *27*, 433–445.
- Freelon, D. (2013). Co-citation map of 9 comm journals, 2003–2013. Retrieved from <http://dfreelon.org/2013/09/05/co-citation-map-of-9-comm-journals-2003-2013>
- Ginther, D. K., & Kahn, S. (2013). Education and academic career out-comes for women of color in science and engineering. In K. Matchett (Ed.), *Seeking solutions: Maximizing American talent by advancing women of color in academia* (pp. 71–92). National Academies Press.
- Gliboff, S. (2018). Sex and the scientific author: M. Vaerting and the Matilda-effect in early twentieth-century Germany. *Gender & History*, *30*(2), 490–510.
- Goldstein, M. L., Morris, S. A., & Yen, G. G. (2005). Group-based Yule model for bipartite author-paper networks. *Physical Review E*, *71*, 026108.
- Gonzalez-Brambila, C. N., Reyes-Gonzalez, L., Veloso, F., & Perez-Angón, M. A. (2016). The scientific impact of developing nations. *PLoS ONE*, *11*, e0151328.
- Goyanes, M. (2020). Editorial boards in communication sciences journals: Plurality or standardization?. *International communication gazette*, *82*(4), 342–364.
- Goyanes, M., & Rodríguez-Gómez, E. F. (2018). ¿ Por qué publicamos? Prevalencia, motivaciones y consecuencias de publicar o perecer. *Profesional de la Información*, *27*(3), 548–558.

- Goyanes, M., & De-Marcos, L. (2020). Academic influence and invisible colleges through editorial board interlocking in communication sciences: a social network analysis of leading journals. *Scientometrics*, 123(2), 791–811.
- Greene, M. (2007). The demise of the lone author. *Nature*, 450, 1165.
- Györfy, B., Herman, P., & Szabó, I. (2020). Research funding: Past performance is a stronger predictor of future scientific output than reviewer scores. *Journal of Informetrics*, 14, 101050.
- Haslam, N., Ban, L., Kaufmann, L., Loughnan, S., Peters, K., Whelan, J., & Wilson, S. (2008). What makes an article influential? predicting impact in social and personality psychology. *Scientometrics*, 76, 169–185.
- Heesen, R., & Romeijn, J. (2019). Epistemic diversity and editor decisions: A statistical Matthew-effect. *Philosophers Imprint*, 19(39), 1–20.
- 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*, 117(9), 4609–4616.
- Iyandemye, J., & Thomas, M. P. (2019). Low income countries have the highest percentages of open access publication: A systematic computational analysis of the biomedical literature. *PLoS ONE*, 14(7), e0220229.
- Jokinen, J., & Pehkonen, J. (2017). Promotions and earnings—gender or merit? evidence from longitudinal personnel data. *Journal of Labor Research*, 38(3), 306–334.
- Kim, M. J. (2001). A bibliometric analysis of physics publications in Korea, 1994–1998. *Scientometrics*, 50, 503–521.
- Knobloch-Westerwick, S., & Glynn, C. J. (2013). The Matilda-effect—role congruity effects on scholarly communication: A citation analysis of communication research and journal of communication articles. *Communication Research*, 40(1), 3–26.
- Knobloch-Westerwick, S., Glynn, C. J., & Huges, M. (2013). The matilda-effect in science communication: An experiment on gender bias in publication quality perceptions and collaboration interest. *Science Communication*, 35(5), 603–625.
- Kurambayev, B., & Freedman, E. (2020). *Publish or Perish? The steep, steep path for Central Asia Journalism and Mass Communication Faculty*. Journalism & Mass Communication Educator: Advance online publication, New York.
- Larivière, V., & Costas, R. (2016). How many is too many? on the relationship between research productivity and impact. *PLoS ONE*, 11(9), e0162709.
- Lauf, E. (2005). National diversity of major international journals in the field of communication. *Journal of Communication*, 55, 19–151.
- Lee, R., & Ellemers, N. (2019). Perceptions of gender inequality in academia. In B. T. Rutjens & M. J. Brandt (Eds.), 2019. London: Belief systems and the perception of reality.
- Lincoln, A. E., Pincus, S., Koster, J. B., & Leboy, P. S. (2012). The Matilda-effect in science: Awards and prizes in the US, 1990s and 2000s. *Social Studies of Science*, 42(2), 307–320.
- Madison, G., & Fahlman, P. (2021). Sex differences in the number of scientific publications and citations when attaining the rank of professor in Sweden. *Studies in Higher Education*, 46(12), 2506–2527.
- Majcher, A. (2002). Gender inequality in German academia and strategies for change. *Global Virtue Ethics Review*, 2(3), 1–35.
- Major, B., & Kaiser, C. R. (2017). Ideology and the maintenance of group inequality. *Group Processes & Intergroup Relations*, 20(5), 582–592.
- Over, R. (1990). The scholarly impact of articles published by men and women in psychological journals. *Scientometrics*, 18, 71–80.
- Plume, A., & Van Weijen, D. (2014). Publish or perish? The rise of the fractional author. *Scholarly Communication*, 3(38). Retrieved from <https://www.researchtrends.com/issue-38-september-2014/publish-or-perish-the-rise-of-the-fractional-author/>
- Santos, L., Cardoso, L., Araújo-Vila, N., & Fraiz-Brea, J. A. (2020). Sustainability perceptions in tourism and hospitality: A mixed-method bibliometric approach. *Sustainability*, 12, 8852.
- Sarsons, H. (2017). Recognition for group work. Working paper. Retrieved at February 15, 2016 via <https://scholar.harvard.edu/sarsons/publications/note-gender-differencesrecognition-group-work>.
- Schröder, M., Lutter, M., & Habicht, I. M. (2021). Publishing, signaling, social capital, and gender: Determinants of becoming a tenured professor in German political science. *PLoS ONE*, 16(1), e0243514.
- Shapiro, F. R. (2000). The most-cited legal scholars. *The Journal of Legal Studies*, 29(S1), 409–426.
- Shen, H. (2013). Inequality quantified: Mind the gender gap. *Nature*, 495(7439), 22–24.
- Stack, S. (2004). Gender, children and research productivity. *Research in higher education*, 45(8), 891–920.
- Stroebe, K., Barreto, M., & Ellemers, N. (2010). Experiencing discrimination: How members of disadvantaged groups can be helped to cope with discrimination. *Social Issues and Policy Review*, 4, 181–213.

- Strumia, A. (2021). Gender issues in fundamental physics: A bibliometric analysis. *Quantitative Science Studies*, 2(1), 225–253.
- Sun, Y., & Genton, M. G. (2011). Functional boxplots. *Journal of Computational and Graphical Statistics*, 20(2), 316–334.
- Valsiner, J. (2006). Dangerous curves in knowledge construction within psychology: Fragmentation of methodology. *Theory & Psychology*, 16(5), 597–612.
- van den Besselaar, P., & Sandström, U. (2015). Gender differences in research performance and its impact on careers: A longitudinal case study. *Scientometrics*, 106, 143–162.
- Wang, D., & Barabási, A. L. (2021). *The Science of Science*. Cambridge University Press.
- Williams, W. M., & Ceci, S. J. (2015). National hiring experiments reveal 2: 1 faculty preference for women on STEM tenure track. *Proceedings of the National Academy of Sciences*, 112(17), 5360–5365.
- Zdeněk, R. (2018). Editorial board self-publishing rates in Czech economic journals. *Science and engineering ethics*, 24(2), 669–682.
- Zetterstrom, R. (2004). The number of authors of scientific publications. *Acta Paediatrica*, 93, 581–582.