



Gender differences among first authors in research focused on the Sustainable Development Goal of *Gender Equality*

Yuanyuan Shang^{1,2} · Gunnar Sivertsen³ · Zhe Cao^{1,2} · Lin Zhang^{1,2,4} 

Received: 30 October 2021 / Accepted: 31 May 2022
© Akadémiai Kiadó, Budapest, Hungary 2022

Abstract

In 2016, the United Nations officially launched the 2030 Agenda for Sustainable Development Goals (SDGs) to address urgent global challenges over the next 15 years. Among the seventeen SDGs, *Gender Equality* (SDG5) is recognized as important for the achievement of the other 16 goals because gender inequality exists across education, employment opportunities, healthcare facilities, life expectancy, family life, and political participation, thereby holding back the capacity of half of the world's population to contribute to solutions to the global challenges. This bibliometric study explores gender balance and differences among first authors within SDG5 oriented research during the first 5 years after the implementation of SDG5 in 2016. Compared with other SDGs, the field of SDG5 produces relatively less scientific publications, which feature a dominance of female first authors. Within the field, male and female first authors focus on partly different topics. Potential readers show more interest in publications by female first authors. This investigation highlights the importance of increasing gender diversity in SDG5-related studies, which is helpful for the achievement of sustainable development.

Keywords Sustainable Development Goals (SDGs) · Gender differences · SDG5 *Gender Equality* · Bibliometric study

✉ Lin Zhang
zhanglin_1117@126.com

¹ Center for Science, Technology & Education Assessment (CSTEa), Wuhan University, Wuhan, China

² Center for Studies of Information Resources, School of Information Management, Wuhan University, Wuhan, China

³ Nordic Institute for Studies in Innovation, Research and Education, Oslo, Norway

⁴ Centre for R&D Monitoring (ECoOM) and Department MSI, KU Leuven, Louvain, Belgium

Introduction

Recognizing the increasing urgency of sustainable development for the world, the United Nations announced the 2030 Agenda for Sustainable Development in September 2015.¹ The agenda lists seventeen Sustainable Development Goals (SDGs) addressing main economic, environmental, and social global challenges. The proposal of the SDGs was immediately understood as encouragement of sustainability research (Leal Filho et al., 2018), recognizing that scientific communities are conducive in translating the SDGs into the priorities and practices of national and local research agendas. As Peter Strohschneider, President of the Deutsche Forschungsgemeinschaft at the time, stated in a German report on how science could contribute, “as a knowledge society, we necessarily rely on scientific research when we try to chart the course towards a sustainable future” (Schmalzbauer & Visbeck, 2016). On this background, the analysis of scientific outputs associated with SDGs is important for the understanding of how research is developing towards contributing to the goals.

As the global community grapples with promoting sustainable development, there is an increasing realization that sustainability and gender equality are deeply interlocked and interdependent (Pandey & Kumar, 2019). Women constitute half of the world’s population. Accordingly, the capacity of women to contribute is vital for the development of society. Gender equality is a major prerequisite for sustainability development. Therefore, among all SDGs, *SDG5 Gender Equality* is significant for achieving the other goals as well. However, gender inequality exists extensively across education, employment opportunities, healthcare facilities, life expectancy, family life, and political participation, thereby hindering the capacity of women to contribute to solutions to the challenges.

Gender equality is a problem within research just as in society in general. Gender differences in academia have been extensively documented (e.g., Larivière et al., 2013; Zhang et al., 2021a, 2021b). In a recent study, Ghiasi et al. (2021) found that, unlike in other SDG-oriented research, female researchers dominate in articles related to *SDG5 Gender Equality*. Our study aims to go deeper into this field of research with a broader array of bibliometric methods to provide a better understanding of current researches related to gender equality and possible suggestions for the achievement of gender equality. To do so, we study the field within a framework of all research oriented towards SDGs in general, especially focusing on first authorship gender gap. The main research questions addressed in this paper are as follows:

- How is research progressing in studies related to *SDG5* compared to SDGs-related studies in general? Does this field of research seem to be given priority?
- What is the gender balance among contributors to research related to *SDG5*? Does the gender balance change?
- Are there gender differences among first authors in research related to *SDG5*?
- Are gender dimensions among first authors in research related to *SDG5* influenced by the human development level of the country in terms of wealth, health, and education?
- Are there gender differences among first authors in the types of research impact of *SDG5*-related articles, citation impact versus readership? What factors influence the research impact of *SDG5*-related articles?

¹ <https://www.un.org/sustainabledevelopment/development-agenda/>.

This paper unfolds as follows. The next section provides a brief background on studies on SDGs-related research and on gender differences in academia. In the section with the title “Data and Methodology”, our strategy for collecting publication data on SDGs-related articles, our procedures for data cleaning, and our methods are presented. In “Results” section, we present our analysis and results of gender differences in SDG5-related articles. The last section reviews and discusses our main findings and their implications, both for research policy and for future work.

Background

Studies on SDGs

Since the SDGs were launched, researchers in all parts of the world have accelerated their scientific explorations to support sustainable development, and research funding organizations in most countries have given extra attention to SDGs-related outcomes. To support overview and policies in this area, bibliometricians have engaged in developing methods by which large datasets representing global research are classified by their relevance for the SDGs. This is not an easy task. The contents and relevance of publications for the SDGs must be identified with multiple methods, such as ontology building, machine learning and hybrid approaches.

Different bibliometric queries have been developed to map publications to SDGs (Pukelis et al., 2020). Different approaches result in various classifications and datasets retrieved from the same general data source, showing that it is important to specify the procedure (Armitage et al., 2020). It seems necessary to accept and encourage multi-angle understandings of SDGs (Rafols et al., 2021). Different bibliometric queries reveal diverse perceptions of the relationship between scientific research and SDGs.

A general observation in these studies is that, at the macro-level, the scientific production related to SDGs has gradually increased in different research areas and regions (Ghiasi et al., 2021). Statistics show that most studies relevant to the SDGs are within the research areas of the Life Sciences & Biomedicine and the Social Sciences. The most predominant SDG among the analyzed research articles is SDG3 *Good Health and Well-being* (Meschede, 2020). However, this observation might be related to the fact that the most used data sources in these studies, Scopus and Web of Science, have a much more extensive coverage of the Health Sciences than they have of the Social Sciences (Aksnes & Sivertsen, 2019).

Studies have also quantified and compared specific SDGs-related research outcomes from countries and institutions across the world, thereby monitoring the global research progress in relation to SDGs (Sweileh, 2020). At the micro-level, relevant studies have taken a single SDG as the research object or mapped the research landscape of SDGs in higher education institutions and other research organizations or explored the potential of SDGs for improving national agendas and policies. These studies are helpful for guiding the research agendas of SDGs (Bautista-Puig et al., 2021; Pineda-Escobar, 2019; Pizzi et al., 2020; Seidman, 2017). Most studies show that SDGs are gradually becoming an important research agenda for both researches performing organizations and research funding organizations.

Studies on gender equality

UN's SDG5 asserts that gender equality is necessary for a sustainable world² because it contributes to higher productivity, good health and well-being, greater economic benefits, more reasonable decisions, and political and economic stability. In this perspective, bibliometric studies focusing on research contributing to the attainment of gender equality are important.

One such study performed by the International Center for the Study of Research (ICSR), which is organized by Elsevier, suggests that relatively few scientific publications contribute to SDG5 *Gender Equality* (Agnew et al., 2020). These publications merely account for about 5 percent of all SDGs-related publications (Ghiasi et al., 2021). It is important, however, to remember that SDG5-related research is mainly provided by the social sciences. They often publish in other languages and publication channels than those covered by Scopus, which is the data source used in the study mentioned above. Aksnes and Sivertsen (2019) find that only 39% of all scientific publications in gender studies from their country (Norway) is covered by Scopus.

The international databases of scientific publications show clear regional differences in SDG5-related studies. High-income countries produce the majority of relevant research (Confraria et al., 2021). Bibliometric studies of gender-related research have also, independently of the SDGs, provided insights into the distribution of countries and institutions in this field of research, and the characteristics of journals publishing the articles (Söderlund & Madison, 2015; Tsay & Li, 2017). Our study aims to provide a deeper understanding of the dynamics of the field in relation to the call for societal contributions to promote SDG5 since 2016 and how the field itself operates from the perspective of gender equity.

Gender differences in academia

Although gender diversity leads to more creative teams and more ground-breaking discoveries (Nielsen et al., 2017), there is still a lack of gender diversity in academia. The difference in scientific performance of scholars of different genders may affect gender representativeness and, therefore, influence gender diversity in the academia. So far, scholars have investigated the differences between male and female researchers in terms of academic productivity, the impact of the publications, research fields and career development. There have been pieces of evidence that the academic productivity of males is higher than that of females in general, but the gap varies in different research areas (Cole & Zuckerman, 1984; Holman et al., 2018; Huang et al., 2020). Nevertheless, this difference may vanish when other variables are taken into account. Productivity differences can be attributed to differences in childcare responsibilities (Kyvik & Teigen, 1996), career trajectories (van den Besselaar & Sandström, 2016), or participation in international collaboration (Aksnes et al., 2019), see also Zhang et al. (2021a). An investigation into French physicists demonstrated that women may appear as productive as men if other determinants of productivity, particularly unequal chances of promotion and other commitments, are controlled (Mairesse & Pezzoni, 2015).

² <https://www.un.org/sustainabledevelopment/gender-equality/>.

As for research impact, there is not a uniform conclusion. Some studies have found that there are fewer citations to articles published by females than those published by males (Bendels et al., 2018; Huang et al., 2020). On the contrary, some studies even found that females are more likely to be cited than males (Nielsen, 2017; Thelwall & Sud, 2020; Thelwall, 2020a, 2020b). Similarly, based on publications by active authors during the period 2014–2018, Elsevier’s gender report (Elsevier, 2020) shows that among the countries studied, the average field-weighted citation impact (FWCI) for men compared to women when assessing all authors, regardless of authorship position, is almost equal in all countries and the EU28. Only when considering the authorship position, the average FWCI of male first authors is higher than that of female first authors (cf. Elsevier, 2020). This phenomenon may be influenced by some factors, such as disciplines, equal opportunities, access to resources and grants. Some studies have also focused on gender differences in other types of impact and found that they can be higher among female researchers. Female researchers seem to have higher visibility among Mendeley readers and in other social media (Bar-Ilan & van der Weijden, 2015; Thelwall, 2018; Zhang et al., 2021a). The gender differences in impacts may be related to differences in the aims of the research. As shown in Zhang et al. (2021a)’s research, male researchers’ studies are mainly aimed at scientific progress, which are more cited. Female researchers’ studies more often engage in studies also aimed at contributing to societal progress, which have more interest among readers (abstract views).

Authorship of publications holds great importance for researchers, which may affect researchers’ academic appointments, promotions, grant funding, and salary support (Kaufmann et al., 2010; Marušić et al., 2011). Prior literatures motivate us to focus on the first-author position. Some studies have shown that the first author is seen as the primary contributor for initial publication conception (Wren et al., 2007). A series of studies have revealed gender differences in the first authorship of academic publications (Rexrode, 2016; Thelwall, 2020a, 2020b; Thelwall & Sud, 2020). In line with previous studies that focus on first authorship, our analyses are based on first author gender only (i.e., female-first versus male-first publications).

Focusing on the gender difference in *SDG5 Gender Equality*, Ghiasi et al. (2021) found in a recent study that female researchers clearly more frequently than male researchers contribute to articles related to *SDG5 Gender Equality*. However, they also found lower citation impact among female researchers in the field and encouraged further studies of these differences. In this study, we further explore the gender differences in *SDG5*-related research because gender participation and contributions from this field of research are important for the achievement of the *SDGs* in general.

Data and methodology

The procedures of data processing are documented in Fig. 1.

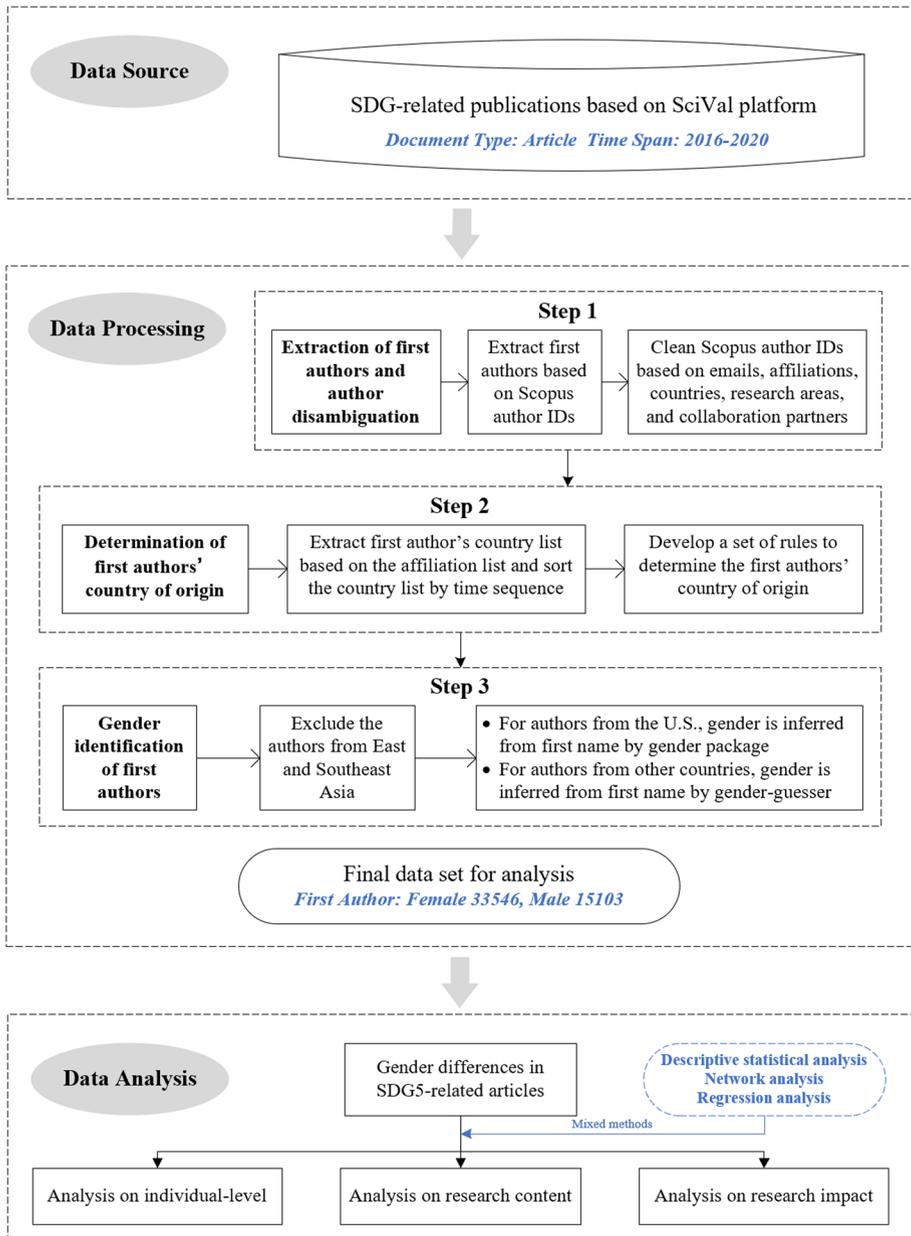


Fig. 1 Procedures of data processing

Data

Data source

There are different methods to classify publications to specific SDGs, including ontology building, supervised machine learning, unsupervised machine learning, and hybrid approaches (Pukelis et al., 2020). Based on these methods, several implementations have been applied, such as the Bergen search approach (Armitage et al., 2020), Elsevier queries (Rivest et al., 2021), AURORA network.³ Different methods can produce different results, which may represent different interpretations of SDGs (Armitage et al., 2020). In fact, there is no single objective “truth” about which research is relevant to reaching SDGs (Rafols et al., 2021). However, bibliometrics should be supportive of a plural landscape to explore various views (Rafols et al., 2021).

Since 2018, Elsevier has generated SDGs search queries to track and demonstrate progress towards SDGs based on Scopus, more than four public versions of queries having been developed so far. The first version in 2019 was created by Elsevier Analytical Services group (Jayabalasingham et al., 2019). Based on this version, Armitage et al. (2020) compared their Bergen search approach with Elsevier queries and found some differences in publications retrieved by the two approaches. Recently, Elsevier updated their search queries by collaborating with AURORA, using machine learning and query improvements, and collected feedback on publications linked to each SDG. The latest version in 2021 implements a new two-stage approach to map publications to the SDGs: (1) Science-Matrix works on new SDGs queries using AURORA Universities Network and Elsevier SDG mapping as references; (2) The machine learning model is designed to improve completeness. As a result, “Elsevier 2021 SDG mapping” captures more publications while keeping precision above 80% (Rivest et al., 2021). Times Higher Education (THE) has applied this approach as part of their 2021 Impact Rankings (Rivest et al., 2021). Therefore, in our study, “Elsevier 2021 SDG mapping” was used to retrieve SDGs-related publications that could be acquired by research areas in SciVal, a research performance assessment tool developed by Elsevier. We selected those indexed as “Article” within the years 2016–2020. The total number of publications for all SDGs is approximately 3.2 million.⁴ Among them, there are 94,863 publications related to SDG5. Data was gathered from June 10, 2021, to June 13, 2021.

Data processing

Extracting researchers’ data from SDG5-related articles

The order in which the authors’ names appear generally reflects each author’s contribution to the article (Baerlocher et al., 2007), with the first author typically playing a lead role in the research and writing process (Larivière et al., 2016). However, it should be noted that some articles may take the “equal contribution” norm, which means that authors use alphabetical sequence to avoid disharmony in collaborating groups (Tschardt et al., 2007). To

³ <https://aurora-universities.eu/resources/educational/sdg-analysis-bibliometrics-relevance/>.

⁴ It should be noted that “Elsevier 2021 SDG mapping” doesn’t cover SDG 17 *Partnerships* for the Goals. Because SDG17 is very difficult to quantify, there is no satisfactory search query to define it.

clarify the percentage of articles listing authors' names in alphabetical order in our dataset, the alphabetical orders of the authors' names were checked. Results showed that the percentage of those listing authors in alphabetical order among articles with no less than three authors was only 5.7%. Therefore, the first authors are considered a proxy for researchers who make major contributions to scientific articles in this study. Based on Scopus Author Identifiers (IDs), 72,273 researchers authored as first authors in SDG5-related articles after duplicating by Scopus author IDs.

It has been shown that Scopus author IDs have high levels of precision and completeness. A study by Paturi and Loktev (2020) showed that the precision and the completeness of Scopus author profiles are 98.3% and 90.6%, respectively. The majority of researchers in Scopus have a single author ID. Nevertheless, due to the high precision of the Scopus author identification system, there are authors with so-called 'split identities' who have more than one Scopus Author ID. To improve the quality of Scopus' individual-level data used in our study, we further conducted author disambiguation procedures manually. We extracted all suspicious author IDs that were likely to be affected by the high precision of Scopus author IDs. An author ID is considered suspicious if the full name is the same as that of another author ID. Firstly, we checked the emails of suspicious author IDs. If the email addresses are the same, they are determined to belong to the same author, and the data of the two suspicious author IDs are merged. Secondly, for suspicious author IDs with the same full name but different emails, then countries, affiliations, research areas and collaboration partners were checked through manual examination. Finally, 148 Scopus author IDs sharing highly similar information with another IDs were merged. After data cleaning of Scopus authors' IDs, there are 72,201 first authors remained in our dataset. In the rest of this paper, we use the term "1st authors" to refer to researchers as defined by author identifiers in Scopus.

Determining a researcher's country of origin

The affiliation list, which is based on the affiliation associated with the author for each publication, is provided in each researcher's Scopus profile. Using Scopus author IDs, we acquired each researcher's affiliation list and extracted corresponding country information. In order to determine the researcher's country of origin, we adopted the method used by Boekhout et al. (2021). In particular, two steps were conducted: (1) For a researcher with affiliations from only one country, the country was marked as the researcher's country of origin. (2) For a researcher with affiliations from more than one country, if the country that a researcher was most often associated with in his/her publications was the same as the country that the researcher was associated with in his/her first publication, we considered this country as the researcher's country of origin. Otherwise, we regarded the evidence as insufficient to determine a single country of origin.

Determining a researcher's gender

Previous studies have shown that a person's first name can be a strong signal of his/her gender (Liu & Ruths, 2013). For each author in our sample, we applied name-based methods to infer a gender. Like other studies, the binary genders are considered and used in our analysis as well (Holman et al., 2018; Santamaría & Mihaljević, 2018). If no gender

information could be inferred for an author, the gender was considered unknown. Our gender assignment process consists of four steps:

Firstly, nonstandard data such as names with no more than two characters were removed.

Secondly, due to the fact that many Asian names are not strongly correlated with gender (Mattauch et al., 2020) and even more gender information gets lost when they have been transliterated in Latin characters, authors from East and Southeast Asia were excluded from our dataset based on the author's country of origin (Huang et al., 2020).⁵

Thirdly, to ensure a relatively high accuracy in the gender classification of the author names and to take the country of origin into the gender assignment process, we used two different name-to-gender inference tools to determine the gender of authors from various countries. For authors from the U.S., gender was inferred from first names by *gender* package⁶ in R, which includes a predictive algorithm and uses historical datasets from the U.S. Social Security Administration, the U.S. Census Bureau, and the North Atlantic Population Project (Blevins & Mullen, 2015). Such data sources are mostly gathered by governments. Researchers have used *gender* package to uncover instances of gender inequity in a range of different areas (Dworkin et al., 2020; Eichmann-Kalwara et al., 2018). For authors from other countries, *gender-guesser* was used to detect gender. The *gender-guesser* package⁷ from Python contains over 45,000 names with gender assignments. The strength of this tool is that gender assignments are checked manually by natives of different countries. *Gender-guesser* has been widely used in scientometric studies and proved to have relatively high accuracy (Adler et al., 2020; Zeina et al., 2020). We gathered the gender of authors inferred by *gender* and *gender-guesser*, respectively. As many as 9533 authors were left with undetermined gender. Our final data set consists of 15,103 males and 33,546 females.

Lastly, we used two independent name-gender data sets to test the accuracy of *gender* and *gender-guesser*. For *gender*, we have manually gathered name-gender data sets from the US universities⁸ and these names are highly likely from the US. The testing data set consists of 1605 names (1011 males, 594 females). When considering only male/female classifications, the accuracy rate of *gender* is 0.9733. For *gender-guesser*, we used a name-gender data set from Wikipedia (Rothe et al., 2015), with names from diversified countries. The testing data set consists of 58,904 names (46,322 males, 12,582 females). When considering only male/female/mostly_male/mostly_female classifications, the accuracy rate of *gender-guesser* is 0.9678.

Data analysis

Individual authors

Groups of the countries based on Human development Index (HDI) Countries of the first authors are classified into four levels (very high, high, medium, and low human development) based on the UNDP's Human development Index (HDI) (United Nations Development Pro-

⁵ In particular, we excluded authors from China (Mainland, Hong Kong, Macau, & Taiwan), Japan, Singapore, Malaysia, South Korea, Indonesia, Thailand, Viet Nam, Philippines, Cambodia, Laos, Myanmar, Mongolia. In total 12,377 (17.1%) authors were excluded from our dataset.

⁶ <https://github.com/lmullen/gender>.

⁷ <https://pypi.python.org/pypi/gender-guesser/>.

⁸ We randomly selected some lists of faculty members from U.S. universities and manually marked the gender information of the names based on faculty members' profile photos.

gram, 2020). The HDI is a summary measure of average achievement in three dimensions of human development—long and healthy life, knowledge, a decent standard of living.

Indicators to reflect a researcher’s focus on SDGs/SDG5-related research Among the articles published by a researcher, the more SDGs-related articles, the more the researcher focuses on SDGs-related topics. Similarly, among the SDGs-related articles published by a researcher, the more SDG5-related articles, the more the researcher focuses on SDG5-related topics. Therefore, we introduce two indicators to reflect a researcher’s research focus on SDGs/SDG5-related research:⁹

- The indicator *Focus on SDGs* refers to the proportion of a researcher’s SDGs-related publications in all his/her publications.
- The indicator *Focus on SDG5* refers to the proportion of a researcher’s SDG5-related publications in all his/her SDGs-related publications.

Research content

All Science Journal Classification of publications Publications are classified using All Science Journal Classification (ASJC) codes from Scopus. These categories comprise different levels of granularity. At the top level, there are five general subject areas: *Life Sciences*, *Physical Sciences*, *Social Sciences*, *Health Sciences*, and *Multidisciplinary*. These five general subject areas are fragmented into 27 “major” subject area classifications, which are further fragmented into 334 “minor” research fields. The content analysis of this study is based on these three levels of ASJC.

Topic cluster and topic cluster prominence Scopus’s data on *Topic Cluster* and *Topic Cluster Prominence* is used to analyze research content at a more granular level. In particular, Scopus takes the entire citation network of Scopus-indexed publications from 1996 and breaks that network into roughly 96,000 topics. When the strength of the citation links between topics reaches a certain threshold, a *Topic Cluster* is formed. There are 1500+ Topic Clusters in total. A publication can only belong to one Topic Cluster.¹⁰

Topic Cluster’s Prominence is an indicator of the momentum/movement or visibility of a particular *Topic Cluster*, which is useful for efforts planning (Klavans & Boyack, 2017). The calculation of *Topic Cluster’s Prominence* combines citations, views, and CiteScore.¹¹ *Topic Clusters prominence* is expressed as percentiles by Scopus. In this study, the *Topic Clusters prominence percentile* (TCPP) is used as a criterion to rank publications.

Research impact

The citation impact and ‘views’ of publications are both considered in this study. *Field-Weighted Citation Impact* (FWCI) in Scopus indicates how the number of citations received by an entity’s publications compares with the average number of citations received

⁹ A researcher’s focus on SDGs/SDG5-related research is based on one’s articles related to SDGs/SDG5 published in 2016–2020.

¹⁰ <https://www.elsevier.com/solutions/scival/features/topic-prominence-in-science>.

¹¹ <https://www.elsevier.com/solutions/scival/features/topic-prominence-in-science>.

by all other similar publications (Elsevier, 2018).¹² Similar to FWCI, *Field-Weighted View Impact* (FWVI) in Scopus indicates how the number of views received by an entity's publications compares with the average number of views received by all other similar publications (Elsevier, 2018). The number of views is the sum of abstract views and clicks on the link to view the full-text at the publisher's website.

Methodology

We used mixed methods to investigate gender differences in the studies associated with SDG5. In particular, descriptive statistical analysis is used to reflect the status quo of gender differences, and network analysis is used to explore the interlinks of research content among SDG5-related articles, and regression analysis is used to investigate the factors affecting the SDG5-related articles' research impact.

To investigate the effect of the gender of 1st author, the HDI of 1st author's country, and the TCPP of the article on the research impact of SDG5-related articles, logistic regression models are adopted in this study. Specifically, the dependent variables of this research are *FWVI* and *FWCI*. As Thelwall and Wilson (2014)'s study shows, a better strategy is to add one to the citations, take their log and then use the general linear model for regression. This strategy is applied to *FWCI* and *FWVI* in this study. The key independent variable is the dummy variable, gender (1 = female, 0 = male). The other independent variables include the HDI of the 1st author's country, and the TCPP of the article. In addition, to eliminate the possible interference of other factors on a publication's impact, the following variables are incorporated as control variables:

- *Collaboration type*, which is the dummy variable (1 means the article has authors from more than one country, and 0 otherwise). Studies have showed that international collaboration can improve citation impact and altmetric indicators (Smith et al., 2014; Wang et al., 2020).
- *N_Authors* refers to the number of authors involved in an article. Studies have proved that the greater the number of authors of a paper, the higher the frequency of citations (Hsu & Huang, 2011). The same pattern has also been seen in social media indicators (Haustein et al., 2015).
- *N_Institutions* refers to the number of institutions involved in an article. Articles involving multiple institutions have been shown to accrue more citations and altmetric impact than single-institution papers (Larivière et al., 2015; Wang et al., 2020).
- *N_Fields* refers to the number of research fields involved in an article.¹³ The relationship between impact and interdisciplinarity of articles is not certain (Chen et al., 2015; Larivière & Gingras, 2010; Wang et al., 2015; Yegros-Yegros et al., 2015). Larivière and Gingras (2010) found that the level of interdisciplinarity of articles and their citation rates are not systematically positively correlated. However, a study by Zhang et al. (2021c) has shown that higher interdisciplinary publications tend to attract more cita-

¹² Similar publications are those publications in the Scopus database that have the same publication year, publication type, and discipline, as represented by the Scopus journal classification system.

¹³ Since a publication may be classified into more than one research field, the full counting approach is used in counting the number of a publication's research fields. Here, the research fields refer to the third level of ASJC in Scopus.

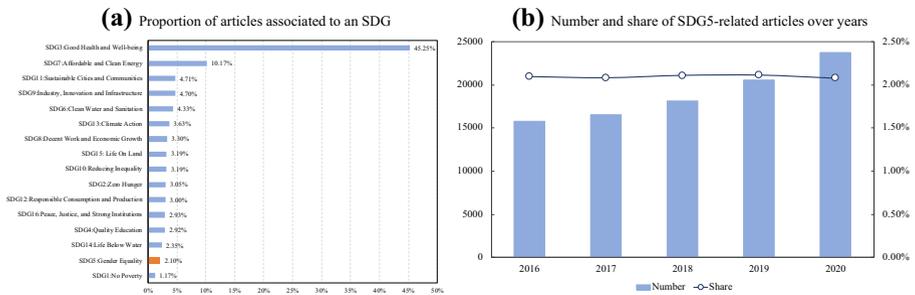


Fig. 2 The distribution of research contributions in each SDG and the number and the proportion of SDG5-related articles over years

tions and have higher PLoS usage (Zhang et al., 2021c). In summary, previous studies have obtained diverging results regarding the relationship between the impact and the interdisciplinarity of articles, probably because different indicators and citation windows were used and different disciplines were focused on.

- *Scimago Journal Rank (SJR)* is proposed that ranks scholarly journals based on citation weighting schemes and eigenvector centrality. The SJR indicator is computed over a journal citation network where the nodes represent the scholarly journals in the Scopus database and the directed connections among the nodes the citation relationships among such journals (González-Pereira et al., 2010). A host of studies have confirmed that publishing papers in journals with higher impact would result in more citations (Tahamtan et al., 2016), which can be easily explained by the positive effect of the vehicles on their contents. Hence, we use SJR, an indicator aims to measure journal prestige, as a control variable as well. We obtained the SJR of all journals from 2016 to 2020 from the SCImago platform.¹⁴ Eventually, 97% of SDG5-related articles can be matched to SJR by ISSN and publication year.
- *Subject* refers to the top level of ASJC (All Science Journal Classification) scheme in Scopus, including Health Sciences, Physical Sciences, Social Sciences & Humanities, Life Sciences and Multidisciplinary. First authorship practices vary between disciplines (West et al., 2013). Citations are also influenced by disciplines (Tahamtan et al., 2016). Dummy variables *Subject1-4* are created to represent this categorical variable.

Results

An overview on SDGs-related articles

As seen in Fig. 2, publications related to SDG5 *Gender equality* accounts for only 2.1% of the 3.2 million SDGs-related articles under study, which is the next to the lowest proportion among all SDGs. The low proportion has been stable between 2016 and 2020 while the number of publications contributing SDG5 has been increasing. In accordance with previous studies (Meschede, 2020) and the statistics reported in *The Power*

¹⁴ <https://www.scimagojr.com/>.

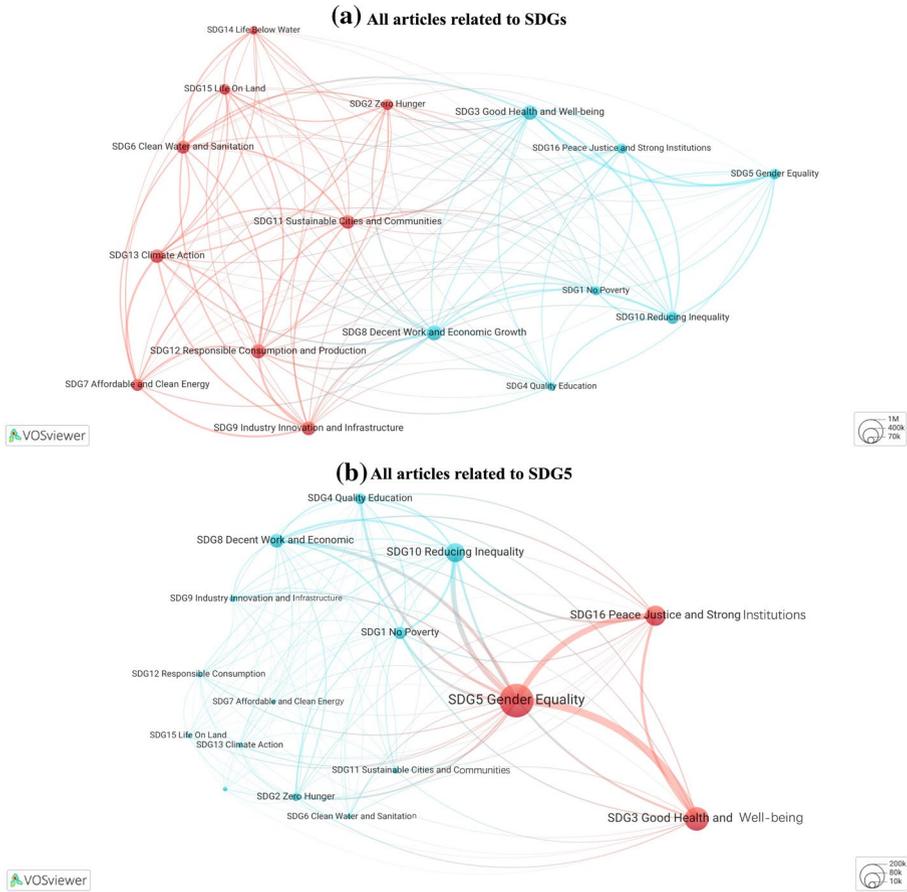


Fig. 3 Co-occurrence networks of SDGs and SDG5 related articles

of Data to Advance the SDGs (Agnew et al., 2020), our results show that SDG3 *Good health and Well-being* is the main focus area of all SDGs-related articles. It should be noted that research that mainly might be based on the social sciences, including research related to Gender Equality, generally have lower proportions in Fig. 2a. Research that may be based on the health sciences, natural sciences and engineering generally have higher proportions which might partly be due to their more comprehensive coverage in Scopus (Aksnes & Sivertsen, 2019).

The SDGs are in many ways dependent on each other, and research can address more than one SDG at the same time. This is visible in our data as the co-occurrence of more than one SDG in the classification of a publication. For a better overview of how the SDGs are related to each other and the position of SDG5 among them, we generated co-occurrence networks among the SDGs in general (Fig. 3a) and for SDG5 in particular (Fig. 3b). The size of a node indicates the total number of articles of the SDG. The thickness of links refers to the strength of co-occurrence between two SDGs measured as the number of articles that address both SDGs. Each color of links and nodes represents a cluster.

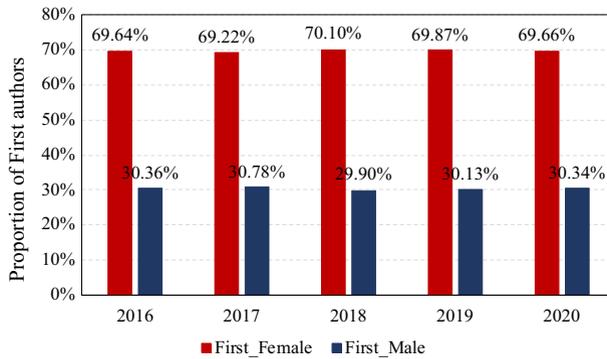


Fig. 4 The proportions of female and male 1st authors of SDG5-related articles over years

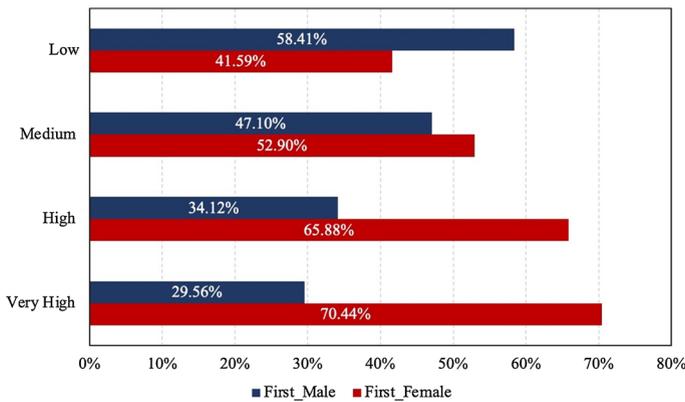


Fig. 5 The proportions of SDG5-related articles of male and female 1st authors by HDI country classification

Among all SDGs-related articles (Fig. 3a), there are two obvious clusters. The first cluster revolves around SDG8 *Decent Work and Economic Growth* and SDG3 *Good health and Well-being*, focusing on the research related to human development issues. The second cluster revolves around SDG12 *Responsible Consumption and Production*, SDG9 *Industry Innovation and Infrastructure*, SDG11 *Sustainable Cities and Communities*, related to the research on global sustainability issues. SDG5 is part of the first cluster focusing on human development, however in a marginal position.

Within articles classified as related to SDG5, Fig. 3b shows the closest relations to SDG3 *Good health and Well-being* and SDG16 *Peace, Justice, and Strong Institutions*. Both are related to human development goals. As an example, SDG5 includes eliminating all forms of discrimination, violence, and harmful practices.¹⁵ Health and well-being are also basic human needs. As seen above in Fig. 3a, the relationships between SDG5 and other SDGs are within the human development cluster. Maslow’s Hierarchy of Needs

¹⁵ <https://www.un.org/sustainabledevelopment/gender-equality/>.

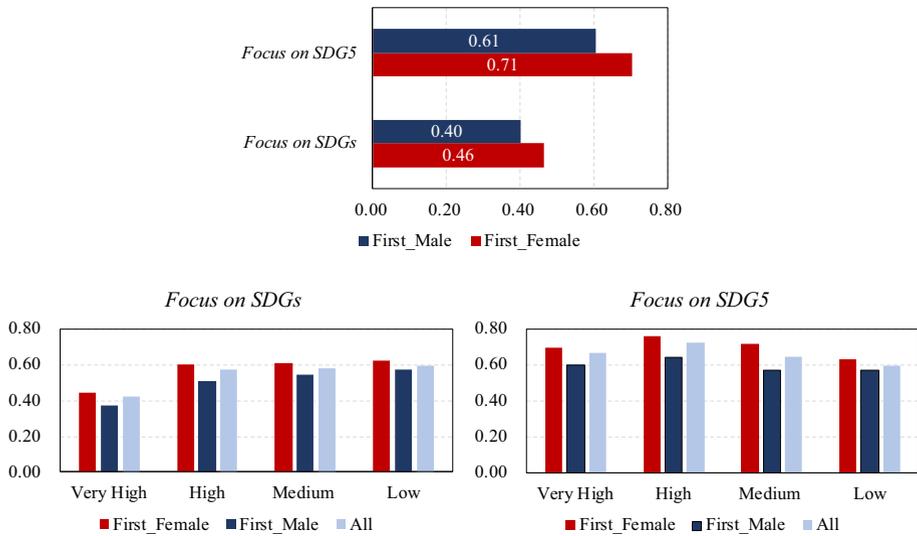


Fig. 6 Average value of *Focus on SDGs* and *Focus on SDG5* by genders and HDI country classification

describes an elevation of human needs from basic and concrete health and safety needs to advanced and abstract concepts such as emotion, reverence, and self-fulfillment (Maslow, 1943). Similarly, the “Four waves of feminism” (Rampton, 2015) shows that the agenda of gender equality historically first focused on fundamentals such as social stability and safety and health, then moved on to economic development, education equality, and other questions of societal equality. This development is shown in both gender studies and in gender equality practice. For instance, four feminism waves respectively promoted female political participation, economic and social equality, as well as full empowerment, which somewhat resemble the co-occurrence network of SDG5-related articles.

Gender differences among first authors of SDG5-related articles

This section explores the characteristics of researchers who conduct the studies related to SDG5, including gender, country of origin, and thematic research profile.

As for gender composition (Fig. 4), females represent a clear majority of 1st authors of SDG5-related articles, with a stable proportion of around 70%. In contrast, many studies have reported the underrepresentation of female first authors in different disciplines (Filardo et al., 2016; Pico et al., 2020; Thelwall & Más-Bleda, 2020). SDG5-related articles represent a domain of research in which female researchers participate much more than elsewhere, while the opposite is true for male researchers. Previous studies have found that women are more aware of gender inequality than men (Popp et al., 2019), which might be conducive for female authors to devote themselves to gender research. However, scientific research requires gender diversity. From this perspective, there should be more male involvement in research focused on gender equality.

The proportion of female 1st authors in SDG5-related research is not the same in all countries. We find that it varies by the countries’ levels of human development. Using the

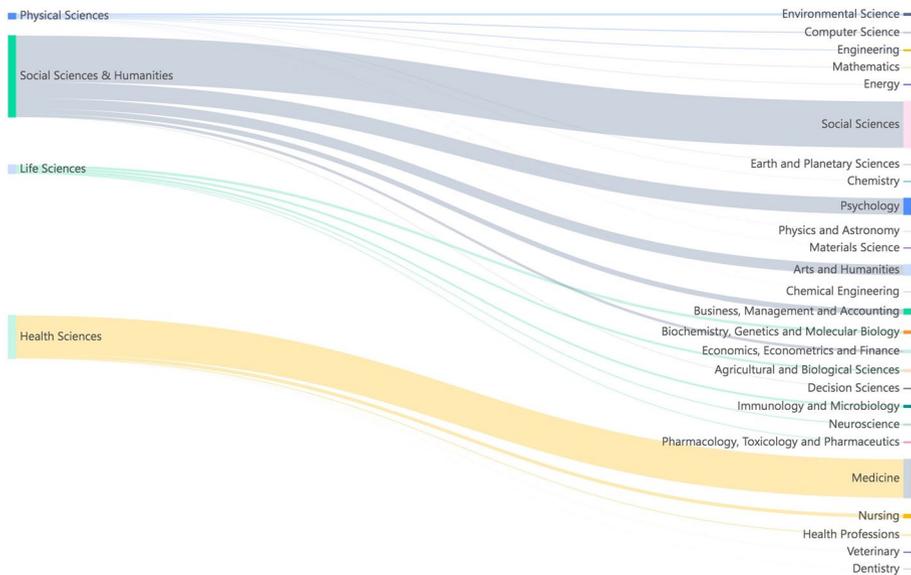


Fig. 7 The distribution of subject areas within major areas of research in SDG5-related articles

Human Development Index (HDI) published by United Nations’ Development Programme (United Nations Development Programme, 2020), we find that high levels of development correspond with high proportions of female researchers in the field (Fig. 5). This result is similar to the results in Ghiasi et al. (2021), where the proportion of female authors in SDG5-related articles is close to 70% in high-income countries. This proportion decreases as the country’s income level decreases. A broader study of the relationship between gender disparities in academia and country-level development found that countries with low HDI have much lower female participation in research in general (Sugimoto et al., 2015). There are many potential explanations for this. One is that in countries with high development levels, females have more opportunities to be educated and are more likely to become researchers. Data from *World Inequality Database on Education* show that females’ mean years of education are longer in high and upper-middle-income countries, whereas males’ mean years of education are longer in low and lower-middle-income countries.¹⁶ Furthermore, in countries with more limited resources, males may be favoured with more resources in academia (Sugimoto et al., 2015). However, during recent decades, women’s political participation has increased to common practice in Latin American countries (Escobar-Lemmon & Taylor-Robinson, 2005), which provides a beneficial social environment to participate in research related to gender equality.

Regarding researchers’ engagement with SDGs-related research, the indicator *Focus on SDGs* indicates the researcher’s degree of attention to SDGs-related topics within the total publication profile, whereas the indicator *Focus on SDG5* shows the degree of attention to SDG5-related topics within the general engagement with SDGs. Figure 6 shows that compared to male 1st authors, female 1st authors are more focused on SDGs in general and concentrated more on SDG5, regardless of the country’s level of human development. Still,

¹⁶ https://www.education-inequalities.org/indicators/eduyears#ageGroup=%22eduyears_2024%22.

Table 1 Gender distribution and gender gap of 1st authors in different subject areas and subject area classifications

Subject Area and Corresponding Subject Area Classification		First Female	First Male	Gender gap
<i>Social Sciences & Humanities</i>		71.66%	28.34%	-43.31%
1	Arts and Humanities**	73.56%	26.44%	-47.12%
2	Social Sciences**	73.48%	26.52%	-46.96%
3	Psychology	71.24%	28.76%	-42.47%
4	Business, Management and Accounting	66.78%	33.22%	-33.55%
5	Decision Sciences	57.85%	42.15%	-15.69%
6	Economics, Econometrics and Finance	57.53%	42.47%	-15.05%
<i>Health Sciences</i>		68.43%	31.57%	-36.86%
1	Nursing**	79.33%	20.67%	-58.66%
2	Health Professions	72.31%	27.69%	-44.62%
3	Medicine	67.54%	32.46%	-35.08%
4	Dentistry	63.64%	36.36%	-27.27%
5	Veterinary	58.55%	41.45%	-17.11%
<i>Physical Sciences</i>		62.62%	37.38%	-25.24%
1	Earth and Planetary Sciences	69.23%	30.77%	-38.46%
2	Engineering	62.52%	37.48%	-25.03%
3	Environmental Science	62.06%	37.94%	-24.11%
4	Computer Science	61.48%	38.52%	-22.96%
5	Energy	61.17%	38.83%	-22.34%
6	Chemistry	53.28%	46.72%	-6.56%
7	Mathematics	52.44%	47.56%	-4.88%
8	Physics and Astronomy*	47.14%	52.86%	5.71%
9	Materials Science*	45.24%	54.76%	9.52%
10	Chemical Engineering*	42.86%	57.14%	14.29%
<i>Multidisciplinary</i>		61.05%	38.95%	-22.09%
<i>Life Sciences</i>		60.63%	39.37%	-21.27%
1	Neuroscience	64.53%	35.47%	-29.07%
2	Pharmacology, Toxicology and Pharmaceutics	59.92%	40.08%	-19.83%
3	Biochemistry, Genetics and Molecular Biology	59.68%	40.32%	-19.36%
4	Agricultural and Biological Sciences	59.41%	40.59%	-18.82%
5	Immunology and Microbiology	58.76%	41.24%	-17.52%

**Indicates the 3 subject area classifications with the highest proportions of female 1st authors

*Indicates the 3 subject area classifications with the highest proportions of male 1st authors

HDI levels are influential. Researchers from countries with lower HDI focus more on studies relevant to SDGs while researchers from countries with higher HDI concentrate relatively more on SDG5-related research. Research related to SDGs that commit to ending poverty and ensuring that all people enjoy peace and prosperity¹⁷ seems to be of comparatively greater importance to low-HDI countries. On the other hand, statistics from World Economic Forum (2019) shows that eight out of the ten countries with the lowest gender gap are very high human development countries (United Nations Development Programme, 2020), indicating a relation between gender equality and interest in research on the same topic.

¹⁷ <https://www.un.org/sustainabledevelopment/development-agenda/>.

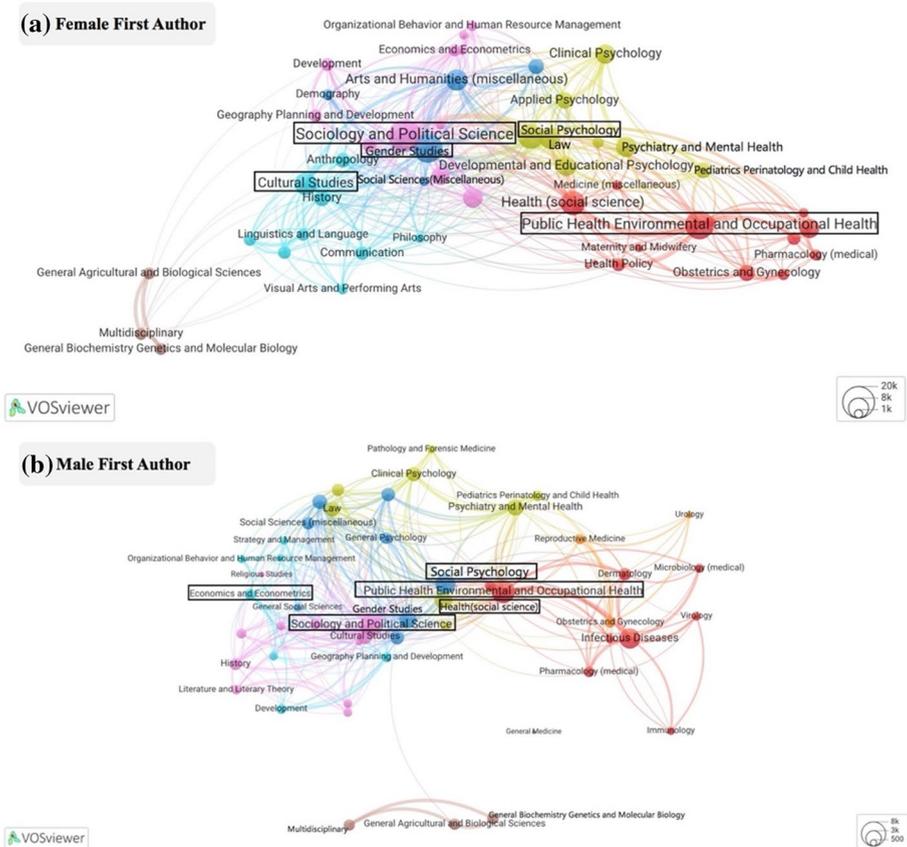


Fig. 8 Co-occurrence networks of top 50 research fields of male and female 1st authors

Gender differences in the research content of SDG5-related articles

To further explore the research content of SDG5-related studies, we first display the distribution of subject areas and subject area classifications (Fig. 7). The subject areas of the SDG5-related articles are mainly distributed in the *Social Sciences & Humanities*, followed by the *Health Sciences*. Within the *Social Sciences & Humanities* as a major area of research, SDG5-related articles mainly occur in subject area classifications of *Social Sciences* and *Psychology*. Within the main area *Health Sciences*, SDG5-related articles mainly occur in the subject area of *Medicine*. Together, these three subject areas seem to be particularly productive in research related to SDG5.

In the following analysis, we present the proportion of male and female 1st authors of SDG5-related articles in each subject area and major area of research. The gender gap is measured as the average proportion of male 1st authors minus that of female 1st authors. As shown in Table 1, although the proportions of female 1st authors are generally higher than those of male 1st authors, there are notable differences among subject areas. Major areas of research with relatively higher proportions of female researchers, such as *Social Sciences & Humanities* and *Health Sciences*, correspondingly have higher proportions

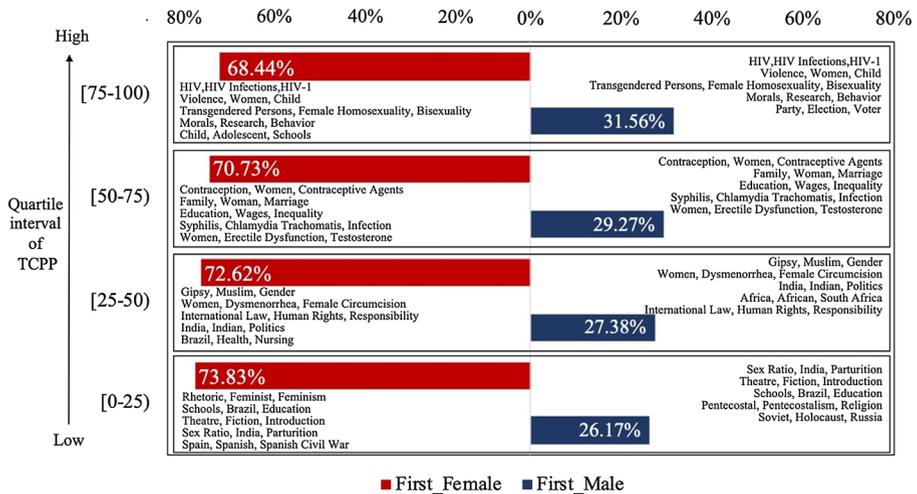


Fig. 9 Proportions of articles and top 5 topic clusters of male and female 1st authors in different groups of topic cluster prominence percentile (TCPP)

of female 1st authors in the subject areas, whereas the proportions of male 1st authors approach 40% in subject areas such as *Physical Sciences* and *Life Sciences*. Likewise, Thelwall and Más-Bleda (2020) demonstrate that gender differences among first authors vary across disciplines: In the *Social Sciences & Humanities* and *Health Sciences*, e.g., in Nursing, Veterinary Sciences, Psychology, and Social Sciences, the proportions of female 1st authors are higher, while in the *Physical Sciences*, e.g., in Physics & Astronomy and Mathematics, the proportions of male 1st authors are higher. Furthermore, the largest gender gaps are found, as expected, in *Art and Humanities*, *Social Sciences* and *Nursing*, all of them known to be female-dominated areas (Elsevier, 2020; Larivière et al., 2013; Thelwall & Mas-Bleda, 2020). By contrast, the gender gaps are smaller in male-dominated areas such as *Chemical Engineering*, *Materials Science*, *Physics and Astronomy* (Elsevier, 2020; Huang et al., 2020; Thelwall & Mas-Bleda, 2020).

Our data allows for further analysis on a third level of more distinct fields of research. In Fig. 8a (female 1st authors) and Fig. 8b (male 1st authors), we visualize co-occurrence networks among the top fifty research fields according to the number of SDG5-related articles. Approaches to gender equality based on *Public Health Environmental and Occupational Health* and *Sociology and Political Science* are important to both genders. Female 1st authors are relatively more active in the cluster named *Cultural Studies* which represents approaches based on *Anthropology*, *History*, *Linguistics and Language* and *Philosophy*. Male 1st authors are relatively more active in the cluster named *Economics and Economics*. Some studies have shown an underrepresentation of women in the economics profession (Auriol et al., 2022; Chevalier, 2020; Larivière et al., 2013) while other studies have shown increasing balance and gender collaboration in the publications (Maddi & Gingras, 2021). It is invisible in our Fig. 8a. In contrast, *Gender Studies* is a more visible cluster for female 1st authors. Among them, this cluster is more related to the humanities and social sciences, including *Arts and Humanities (miscellaneous)*, *Demography*, *History and Philosophy of Science*, while among male 1st authors (Fig. 8b) *Gender Studies* are more related to psychology, such as *Social Psychology*, *General Psychology*. These observations

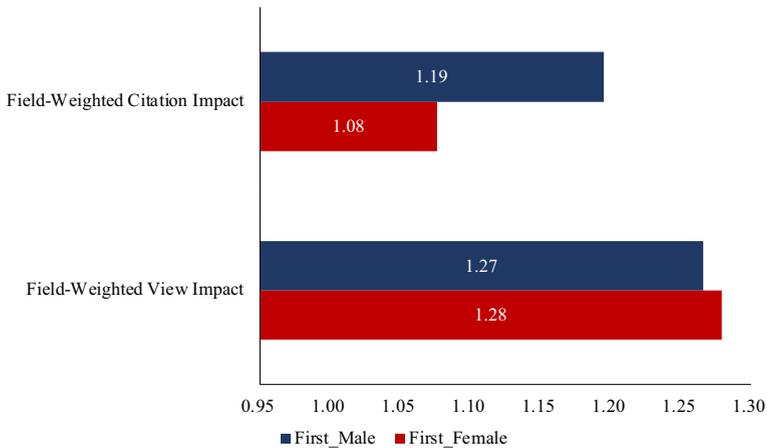


Fig. 10 Mean values of *FWCI* and *FWVI* of articles authored by male and female 1st authors

are also consistent with general gender differences in corresponding research fields. Studies have shown that disciplines from the social sciences have a relatively larger proportion of female authors (Larivière et al., 2013), while males, on average, dominate more in psychology publications (König et al., 2015).

Using the methodology and indicators of *Topic Cluster’s Prominence* as introduced in the section of *Data analysis* above, we will now analyze research content at an even more granular level. Topic Clusters are found by citation networks while the Prominence of a cluster is based on citation impact and readers’ views of articles in Scopus. In this study, the *Topic Clusters prominence percentile* (TCPP) is used as a criterion to rank the publications within a cluster according to impact. As shown in Fig. 9, we find that the lower the topic cluster prominence percentile, the higher the proportion of females contributing to the topic cluster. SDG5-related topic clusters with less prominence are primarily promoted by women. The topic clusters themselves are more similar between males and females among those with comparatively higher prominence. We only see that females take more interest in *Child, Adolescent, Schools* and men take more interest in political issues including *Party, Election, Voter*. As for the topic clusters with relatively lower prominence, research topics of both genders feature regional characteristics, focusing on regions including Africa, India, Brazil, Russia, and Spain. Part of the gender differences found in this analysis might be related to those found above. Topic clusters based on the health sciences seem to be more prominent than those based on the social sciences, i.e., in areas of research with higher proportions of male researchers.

Gender differences in research impact of SDG5-related articles

Several of studies have found that articles by female researchers tend to be less cited than articles by male researchers, but some studies have not confirmed this. A recent overview of the literature on gender and impact is given by Zhang et al. (2021a) where one explanation for the difference is also relevant here: female researchers more often engage in

Table 2 Descriptive statistics and correlations

Variable	Mean	SD	Max	Variance	Min	1	2	3	4	5	6	7	8
1. FWVI	1.30	1.98	291.66	3.91	0.00								
2. FWCI	1.13	2.68	313.37	7.18	0.00	0.520***							
3. N_Authors	4.31	10.34	1217.00	106.83	1.00	0.433***	0.235***						
4. N_Institutions	2.13	3.85	512.00	14.83	1.00	0.589***	0.334***	0.769***					
5. N_Fields	2.20	1.12	11.00	1.25	1.00	0.016***	0.017***	-0.002	-0.003				
6. SJR	1.01	1.21	40.95	1.47	0.10	0.119***	0.273***	0.187***	0.197***	0.029***			
7. TCPP	72.22	25.17	99.80	633.78	0.20	0.037***	0.045***	0.114***	0.097***	0.046***	0.173***		
8. HDI	0.89	0.09	0.96	0.01	0.40	0.026***	0.072***	-0.006	0.012***	0.060***	0.120***	0.006	

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3 Regression analysis results (dependent variables are *FWVI* and *FWCI*)

Ivs	ln (FWVI)			ln (FWCI)		
	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)	Model (6)
ln(N_Authors)	0.039*** (0.00)	0.041*** (0.00)	0.031*** (0.00)	0.028*** (0.01)	0.033*** (0.01)	0.023*** (0.01)
ln(N_Institutions)	0.024*** (0.01)	0.019*** (0.01)	0.023*** (0.01)	0.056*** (0.01)	0.048*** (0.01)	0.056*** (0.01)
ln(N_Fields)	0.045*** (0.01)	0.043*** (0.01)	0.044*** (0.01)	0.036*** (0.01)	0.032*** (0.01)	0.035*** (0.01)
Collaboration type	0.015*** (0.00)	0.018*** (0.00)	0.013*** (0.00)	0.000 (0.01)	0.009 (0.01)	0.000 (0.01)
ln (SJR)	0.105*** (0.01)	0.096*** (0.01)	0.097*** (0.01)	0.564*** (0.01)	0.545*** (0.01)	0.559*** (0.01)
Subject1	-0.018*** (0.00)	-0.019*** (0.00)	-0.023*** (0.00)	-0.070*** (0.01)	-0.071*** (0.01)	-0.073*** (0.01)
Subject2	0.073*** (0.01)	0.073*** (0.01)	0.068*** (0.01)	-0.010 (0.01)	-0.006 (0.01)	-0.012 (0.01)
Subject3	0.028*** (0.01)	0.024*** (0.01)	0.028*** (0.01)	0.066*** (0.01)	0.057*** (0.01)	0.065*** (0.01)
Subject4	-0.092*** (0.01)	-0.092*** (0.01)	-0.096*** (0.01)	-0.142*** (0.01)	-0.136*** (0.01)	-0.142*** (0.01)
Gender (Reference male)	0.025*** (0.00)			-0.002 (0.00)		
HDI		0.165*** (0.02)			0.359*** (0.02)	
TCPP			0.001*** (0.00)			0.0005*** (0.00)
Constant	0.517*** (0.01)	0.397*** (0.02)	0.502*** (0.01)	0.083*** (0.01)	-0.218*** (0.02)	0.059*** (0.01)
Prob > F	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
R-squared	0.0262	0.0265	0.0272	0.1642	0.1671	0.1646

① * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$; ② $N = 61,557$; ③ Robust standard errors in parentheses.

research aimed at societal progress. Their publications are less cited, but more read. We find the same in this study.

So far, little is known about gender differences in research impact in studies related to gender equality. We study research impact as expressed both by citations from the scientific literature and as the interest in the publication among potential readers. Citation impact is measured within Scopus using the indicator *Field-Weighted Citation Impact* (FWCI) as presented in section of *Data analysis*. Interest among readers is measured by the indicator *Field-Weighted View Impact* (FWVI) as presented in the same section. The view count is the sum of abstract views and clicks on the link to view the full text at the publisher’s website.

Figure 10 presents the mean value of *FWVI* and *FWCI* of male/female 1st authors in SDG5-related articles. Female 1st authors have a relatively higher mean value of *FWVI*, whereas male 1st authors have a relatively higher mean value of *FWCI*. The Mann–Whitney

U test is performed on view and citation impact for different gender groups. If the $p < 0.05$, it indicates significant gender differences. The Mann–Whitney U tests show that there are significant differences between male and female 1st authors ($FWCI: p = 0.0483$, $FWVI: p = 0.000$). Results suggests that the SDG5-related articles by male 1st authors have higher citation impact while those by female 1st authors have more interest among readers, confirming the more general results in studies by Thelwall (2018) and Zhang et al. (2021a). Our general finding above that female 1st authors more often engage in SDGs-related research is accordance with the observation in the latter study that female researchers more often value and engage in research aimed at societal progress and thereby find more interest among potential readers.

To better understand the observed gender differences in the research impact of SDG5-related articles, we use Ordinary Least Squares (OLS) regressions in STATA 15.1 to detect the differences of research impact after other variables are added to the models. It should be noted that because of the discrete lognormal distribution of data, we use the natural logarithms of $FWCI$, $FWVI$, $N_Authors$, $N_Institutions$, N_fields , SJR as variables. White's test shows the heteroskedasticity is considerable, so we use the `vce(robust)` option to substitute a robust variance matrix calculation for the conventional calculation.

Descriptive statistics and correlations of the examined variables are presented in Table 2. The test results in Table 2 found a significant correlation between almost all variables, which might lead to a multicollinearity problem to a certain extent. Therefore, the multicollinearity test was carried out in this research. Using $VIF > 5$ as cut-off value (Berk, 2004; Liao & Yen, 2012), all predictors and covariates had VIF values below this threshold, indicating no multicollinearity problem between the variables in the regression model.

The statistical tests thereby support the interpretation of Fig. 10 above. The regression analysis results in Table 3 below show that female 1st authors create a significant positive impact on view impact (Model (1)), whereas their impact on citation impact may not be significant (Model (4)). Our findings confirm the results of previous studies (Ledin et al., 2007; Penas & Willett, 2006): Female first authors publish articles with a higher view impact, and the reason might be that they more often value and engage in research aimed at contributing to societal progress (Zhang et al. 2021a).

Furthermore, Model (2) and Model (5) show the influence of the human development level of 1st author's country on view and citation impact. The results show that the higher value of the HDI of 1st author's country, the greater impact of both impacts. As for the topic cluster prominence of SDG5-related articles, the $TCPP$ of SDG5-related articles is positively and significantly ($p < 0.05$) related to $FWVI$ and $FWCI$ (Model (3) and Model (6)).

Discussions and conclusions

This study was piqued by an aim to underpin current global efforts to promote gender diversity in the studies related to SDG5, which matters for the achievement of gender equality in research and in society. We draw four main conclusions from our analysis:

Firstly, SDG5-related articles account for a small proportion of all SDGs-related articles. The share of articles relevant to SDG5 has not increased during the first five years after the launch of SDGs. Furthermore, SDG5-related articles are in a relatively peripheral position in the network of SDGs-related articles. Although these results may partly reflect a limited coverage of the social sciences in Scopus, it seems that studies associated with

SDG5 have not received widespread attention in academia. As stated in the introduction, females play a critical role in realizing SDGs in general, and SDG5 profoundly impacts the realization of every other SDG (Pandey & Kumar, 2019). To take two examples, for SDG1 *No Poverty*, the elimination of gender bias may improve women's productivity and earnings, which might reduce poverty and accelerate economic growth (Morrison et al., 2007). From the perspective of SDG2 *Zero Hunger*, in nearly two-thirds of all countries, women are more likely than men to report food insecurity.¹⁸ Females are faced with different forms of discrimination within the family, which may deny their access to basic facilities, including an essential diet. Therefore, gender equality is important to end all forms of malnutrition.

Secondly, female 1st authors dominate studies associated with SDG5, which reflect a gender imbalance which is the opposite in most other fields or research. Moreover, compared with males, females are more concerned with SDGs-related studies in general and concentrate more on SDG5-related research. This may indicate that females are more perceptive and willing to research for gender equality, highlighting the need for gender diversity in SDG5-related research. Furthermore, female 1st authors are highly represented in scientific activities related to SDG5 in countries with very high, high, and medium levels of human development (HDI), while there are relatively more male 1st authors in countries with low HDI. The level of development, the social environment, and the level of education may influence the degrees of gender participation in SDG5-related studies in different countries.

Thirdly, as for the research content, the researches in the SDG5-related articles are mainly based on the *Social Sciences & Humanities* and the *Health Sciences*. In general, these main areas of research, as well as the subject areas of the publications, in general have relatively higher proportions of female researchers which corresponds to the higher proportion of female 1st authors in SDG5-related articles. Within the main areas of research, male and female researchers partly focus on different fields of research and topics when performing research related to SDG5. Male researchers focus more on traditionally male-dominated fields and the more prominent topics with regard to impact. Also, for this reason, it is important to enhance gender diversity in SDG5-related studies to promote broader research on SDG5.

Lastly, our findings show that compared to male 1st authors, SDG5-related publications by female 1st authors have relatively lower citation impact and slightly higher interest among readers (measured as views). This is also evidenced by the regression analysis, with female 1st authors have a positive impact on the view indicator. Similar results and a general framework for understanding this finding have recently been presented in Zhang et al. (2021a): Female researchers value and engage relatively more in research aimed at societal progress. Their publications gain more interest among readers. Our study's finding that female researchers engage relatively more in SDGs-related research is in accordance with this general framework.

The limitations of this study are, firstly, that the analysis is mainly focused on the first authors. Many of the scientific publications under study were probably published in collaboration by female and male researchers. Future research could extend the analysis to all authors and examine the gender differences through a more comprehensive lens. Secondly, as already commented, limitations in the coverage of the social sciences in Scopus may be a factor influencing the results.

¹⁸ <https://www.unwomen.org/en/news/in-focus/women-and-the-sdgs/sdg-2-zero-hunger>.

Despite these limitations, it seems clear from our findings that gender differences among first authors in SDG5-related research do exist in productivity, research impact and research topic. There are different topical focuses and impact advantages in scientific output of male and female first authors. This may imply a need to increase gender diversity in these studies, which can make it possible to give full play to the strengths of authors of different genders. As previous studies show, gender diversity can stimulate better reflections that may lead to more original results (Maddi & Gingras, 2021). Research related to *Gender Equality* is needed for the achievement of SDGs in general and may improve fairness, equity, and diversity in the progress of sustainable development.

Acknowledgements The authors would like to acknowledge support from the National Natural Science Foundation of China (Grant Nos. 71974150, 72004169, 72074029), the National Laboratory Centre for Library and Information Science at Wuhan University, and the Research Council of Norway, Grant 256223, FORINNPOL.

Declarations

Conflict of interest The corresponding author (Lin Zhang) is the Co-Editor-in-Chief of *Scientometrics*.

References

- Adler, E., Hobbs, A., Dhaliwal, G., & Babik, J. M. (2020). Gender differences in authorship of clinical problem-solving articles. *Journal of Hospital Medicine*, 15(8), 475–478.
- Agnew, K., Francescon, D., Martin, R., Rhannam, M., & Schemm, Y. (2020). *The power of data to advance the SDGs*. Retrieved from https://www.elsevier.com/_data/assets/pdf_file/0004/1058179/Elsevier-SDG-Report-2020.pdf. Accessed 9 Sept 2021.
- Aksnes, D. W., Piro, F. N., & Rørstad, K. (2019). Gender gaps in international research collaboration: A bibliometric approach. *Scientometrics*, 120, 747–774.
- Aksnes, D. W., & Sivertsen, G. (2019). A criteria-based assessment of the coverage of scopus and web of science. *Journal of Data and Information Science*, 4(1), 1–21.
- Armitage, C. S., Lorenz, M., & Mikki, S. (2020). Mapping scholarly publications related to the sustainable development goals: Do independent bibliometric approaches get the same results? *Quantitative Science Studies*, 1(3), 1092–1108.
- Auriol, E., Friebe, G., Weinberger, A., & Wilhelm, S. (2022). Underrepresentation of women in the economics profession more pronounced in the United States compared to heterogeneous Europe. *Proceedings of the National Academy of Sciences of the United States of America*, 119(16), e2118853119.
- Baerlocher, M. O., Newton, M., Gautam, T., Tomlinson, G., & Detsky, A. S. (2007). The meaning of author order in medical research. *Journal of Investigative Medicine*, 55(4), 174–180.
- Bar-Ilan, J., & van der Weiden, I. (2015). Altmetric gender bias? An exploratory study. *International Journal of Computer Science: Theory and Application*, 4(1), 16–22.
- Bautista-Puig, N., Aleixo, A. M., Leal, S., Azeiteiro, U., & Costas, R. (2021). Unveiling the research landscape of sustainable development goals and their inclusion in Higher Education Institutions and Research Centers: Major trends in 2000–2017. *Frontiers in Sustainability*, 2, 620743.
- Bendels, M. H., Müller, R., Brueggemann, D., & Groneberg, D. A. (2018). Gender disparities in high-quality research revealed by Nature Index journals. *PLoS ONE*, 13(1), e0189136.
- Berk, R. A. (2004). *Regression analysis: A constructive critique*. Sage.
- Blevins, C., & Mullen, L. (2015). Jane, John ... Leslie? A historical method for algorithmic gender prediction. *Digital Humanities Quarterly*, 9(3), 000223.
- Boekhout, H., van der Weijden, I., & Waltman, L. (2021). Gender differences in scientific careers: A large-scale bibliometric analysis. Preprint retrieved from <https://arxiv.org/abs/2106.12624>
- Chen, S., Arsenaault, C., & Larivière, V. (2015). Are top-cited papers more interdisciplinary? *Journal of Informetrics*, 9(4), 1034–1046.
- Chevalier, J. (2020). *The 2020 Report of the Committee on the Status of Women in the Economics Profession*. Committee on the Status of Women in the Economics Profession, American Economic Association.

- Cole, J. R., & Zuckerman, H. (1984). The productivity puzzle: Persistence and change in patterns of publication of men and women scientists. *Advances in Motivation and Achievement*, 2, 217–256.
- Confraria, H., Noyons, E., & Ciarli, T. (2021). Countries' research priorities in relation to the Sustainable Development Goals. In W. Glänzel, S. Heeffer, P. S. Chi, & R. Rousseau (Eds.), *Proceedings of the 18th International Conference on Scientometrics and Informetrics* (pp. 281–291)
- Dworkin, J. D., Linn, K. A., Teich, E. G., Zurn, P., Shinohara, R. T., & Bassett, D. S. (2020). The extent and drivers of gender imbalance in neuroscience reference lists. *Nature Neuroscience*, 23(8), 918–926.
- Eichmann-Kalwara, N., Jorgensen, J., & Weingart, S. B. (2018). Representation at Digital Humanities Conferences (2000–2015). *Bodies of Information: Intersectional feminism and digital humanities* (pp. 72–92).
- Elsevier. (2018). *Research metrics guidebook*. Retrieved from <https://www.elsevier.com/research-intelligence/resource-library/research-metrics-guidebook>. Accessed 9 Sept 2021.
- Elsevier. (2020). *The researcher journey through a gender lens: An examination of research participation, career progression and perceptions across the globe*. Retrieved from <https://www.elsevier.com/connect/gender-report>. Accessed 9 Sept 2021.
- Escobar-Lemmon, M., & Taylor-Robinson, M. M. (2005). Women ministers in Latin American government: When, where, and why? *American Journal of Political Science*, 49(4), 829–844.
- Filardo, G., da Graca, B., Sass, D. M., Pollock, B. D., Smith, E. B., & Martinez, M. A. (2016). Trends and comparison of female first authorship in high impact medical journals: Observational study (1994–2014). *BMJ*, 352, 847.
- Ghiasi, G., Harsh, M., Tajmel, T., & Larivière, V. (2021). Where international development and gender equality meet in science: a bibliometric analysis. In W. Glänzel, S. Heeffer, P. S. Chi, & R. Rousseau (Eds.), *Proceedings of the 18th International Conference on Scientometrics and Informetrics* (pp. 447–452).
- González-Pereira, B., Guerrero-Bote, V. P., & Moya-Anegón, F. (2010). A new approach to the metric of journals' scientific prestige: The SJR indicator. *Journal of Informetrics*, 4(3), 379–391.
- Haustein, S., Costas, R., & Larivière, V. (2015). Characterizing social media metrics of scholarly papers: The effect of document properties and collaboration patterns. *PLoS ONE*, 10(3), e0120495.
- Holman, L., Stuart-Fox, D., & Hauser, C. E. (2018). The gender gap in science: How long until women are equally represented? *PLoS Biology*, 16(4), e2004956.
- Hsu, J.-W., & Huang, D.-W. (2011). Correlation between impact and collaboration. *Scientometrics*, 86(2), 317–324.
- 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.
- Jayabalasingham, B., Boverhof, R., Agnew, K., & Klein, L. (2019). *Identifying research supporting the United Nations Sustainable Development Goals*. Mendeley Data, V1. <https://doi.org/10.17632/87txkw7khs.1>
- Kaufmann, P., Annis, C., Griggs, R. C., Muscle Study Group Executive Committee. (2010). The authorship lottery: An impediment to research collaboration? *Annals of Neurology*, 68(6), 782–786.
- Klavans, R., & Boyack, K. W. (2017). Research portfolio analysis and topic prominence. *Journal of Informetrics*, 11(4), 1158–1174.
- König, C. J., Fell, C. B., Kellnhöfer, L., & Schui, G. (2015). Are there gender differences among researchers from industrial/organizational psychology? *Scientometrics*, 105(3), 1931–1952.
- Kyvik, S., & Teigen, M. (1996). Child care, research collaboration, and gender differences in scientific productivity. *Science, Technology, & Human Values*, 21(1), 54–71.
- Larivière, V., Desrochers, N., Macaluso, B., Mongeon, P., Paul-Hus, A., & Sugimoto, C. R. (2016). Contributorship and division of labor in knowledge production. *Social Studies of Science*, 46(3), 417–435.
- Larivière, V., & Gingras, Y. (2010). On the relationship between interdisciplinarity and scientific impact. *Journal of the American Society for Information Science and Technology*, 61(1), 126–131.
- Larivière, V., Gingras, Y., Sugimoto, C. R., & Tsou, A. (2015). Team size matters: Collaboration and scientific impact since 1900. *Journal of the Association for Information Science and Technology*, 66(7), 1323–1332.
- Larivière, V., Ni, C., Gingras, Y., Cronin, B., & Sugimoto, C. R. (2013). Bibliometrics: Global gender disparities in science. *Nature News*, 504(7479), 211.
- Leal Filho, W., Azeiteiro, U., Alves, F., Pace, P., Mifsud, M., Brandli, L., et al. (2018). Reinvigorating the sustainable development research agenda: The role of the sustainable development goals (SDG). *International Journal of Sustainable Development & World Ecology*, 25(2), 131–142.
- Ledin, A., Bornmann, L., Gannon, F., & Wallon, G. (2007). A persistent problem—Traditional gender roles hold back female scientists. *EMBO Reports*, 8(11), 982–987.

- Liao, C. H., & Yen, H. R. (2012). Quantifying the degree of research collaboration: A comparative study of collaborative measures. *Journal of Informetrics*, 6(1), 27–33.
- Liu, W., & Ruths, D. (2013). What's in a name? using first names as features for gender inference in twitter. In *Analyzing microtext: 2013 AAAI spring symposium*.
- Maddi, A., & Gingras, Y. (2021). Gender diversity in research teams and citation impact in Economics and Management. *Journal of Economic Surveys*, 35(5), 1381–1404.
- Mairesse, J., & Pezzoni, M. (2015). Does gender affect scientific productivity: A critical review of the empirical evidence and a panel data econometric analysis for french physicists. *Revue Économique*, 66(1), 65–113.
- Marušić, A., Bošnjak, L., & Jerončić, A. (2011). A systematic review of research on the meaning, ethics and practices of authorship across scholarly disciplines. *PLoS ONE*, 6(9), e23477.
- Maslow, A. H. (1943). A theory of human motivation. *Psychological Review*, 50(4), 370–396.
- Mattauch, S., Lohmann, K., Hannig, F., Lohmann, D., & Teich, J. (2020). A bibliometric approach for detecting the gender gap in computer science. *Communications of the ACM*, 63(5), 74–80.
- Meschede, C. (2020). The sustainable development goals in scientific literature: A bibliometric overview at the meta-level. *Sustainability*, 12(11), 4461.
- Morrison, A., Raju, D., & Sinha, N. (2007). Gender equality, poverty and economic growth. *Policy Research Working Paper, No. 4349*. Retrieved from <https://openknowledge.worldbank.org/handle/10986/7321>. Accessed 15 Sept 2021.
- Nielsen, M. W. (2017). Gender and citation impact in management research. *Journal of Informetrics*, 11(4), 1213–1228.
- Nielsen, M. W., Alegria, S., Börjeson, L., Etzkowitz, H., Falk-Krzesinski, H. J., Joshi, A., et al. (2017). Opinion: Gender diversity leads to better science. *PNAS*, 114(8), 1740–1742.
- Pandey, U. C., & Kumar, C. (2019). The relationship of SDG5 to other goals. In U. C. Pandey & C. Kumar (Eds.), *SDG5—gender equality and empowerment of women and girls (concise guides to the united nations sustainable development goals)* (pp. 103–120). Emerald Publishing Limited.
- Paturi, M., & Loktev, A. (2020). The best gets better: Scopus data quality measured. In *Pure International Conference*. Retrieved from <https://brighttalk.com/webcast/13819/456949/prcn2020-the-best-gets-better-scopus-data-quality-measured>. Accessed 15 Sept 2021.
- Penas, C. S., & Willett, P. (2006). Gender differences in publication and citation counts in librarianship and information science research. *Journal of Information Science*, 32(5), 480–485.
- Pico, T., Bierman, P. R., Doyle, K., & Richardson, S. (2020). First Authorship Gender Gap in the Geosciences. *Earth and Space Science*, 7(8), e2020EA001203.
- Pineda-Escobar, M. A. (2019). Moving the 2030 agenda forward: SDG implementation in Colombia. *Corporate Governance: The International Journal of Business in Society*, 19(1), 176–188.
- Pizzi, S., Caputo, A., Corvino, A., & Venturelli, A. (2020). Management research and the UN sustainable development goals (SDGs): A bibliometric investigation and systematic review. *Journal of Cleaner Production*, 276, 124033.
- Popp, A. L., Lutz, S. R., Khatami, S., Van Emmerik, T. H., & Knoben, W. J. (2019). A global survey on the perceptions and impacts of gender inequality in the Earth and space sciences. *Earth and Space Science*, 6(8), 1460–1468.
- Pukelis, L., Puig, N. B., Skrynik, M., & Stanciauskas, V. (2020). OSDG—Open-Source approach to classify text data by UN sustainable development goals (SDGs). Preprint retrieved from <https://arxiv.org/abs/2005.14569>.
- Rafols, I., Noyons, E., Confraria, H., & Ciarli, T. (2021). Visualising plural mappings of science for Sustainable Development Goals (SDGs). In: W. Glänzel, S. Heeffer, P. S. Chi, & R. Rousseau (Eds.), *Proceedings of the 18th International Conference on Scientometrics and Informetrics* (pp. 949–954).
- Rampton, M. (2015). Four waves of feminism. *Pacific University Oregon*, 25. Retrieved from <https://www.pacificu.edu/magazine/four-waves-feminism>. Accessed 15 Sept 2021.
- Rexrode, K. M. (2016). The gender gap in first authorship of research papers. *BMJ*, 352, 1130.
- Rivest, M., Kashnitsky, Y., Bédard-Vallée, A., Campbell, D., Khayat, P., Labrosse, I., et al. (2021). *Improving the Scopus and Aurora queries to identify research that supports the United Nations Sustainable Development Goals (SDGs) 2021, Mendeley Data*, V2. <https://doi.org/10.17632/9sxdykm8s4.2>
- Rothe, R., Timofte, R., & Van Gool, L. (2015). Dex: Deep expectation of apparent age from a single image. In *2015 IEEE International Conference on Computer Vision Workshop (ICCVW)* (pp. 10–15). <https://doi.org/10.1109/ICCVW.2015.41>.
- Santamaría, L., & Mihaljević, H. (2018). Comparison and benchmark of name-to-gender inference services. *PeerJ Computer Science*, 4, e156.
- Schmalzbauer, B., & Visbeck, M. (2016). *The contribution of science in implementing the Sustainable Development Goals*. German Committee Future Earth.

- Seidman, G. (2017). Does SDG 3 have an adequate theory of change for improving health systems performance? *Journal of Global Health*, 7(1), 010302.
- Smith, M. J., Weinberger, C., Bruna, E. M., & Allesina, S. (2014). The scientific impact of nations: Journal placement and citation performance. *PLoS ONE*, 9(10), e109195.
- Söderlund, T., & Madison, G. (2015). Characteristics of gender studies publications: A bibliometric analysis based on a Swedish population database. *Scientometrics*, 105(3), 1347–1387.
- Sugimoto, C. R., Ni, C., & Larivière, V. (2015). On the relationship between gender disparities in scholarly communication and country-level development indicators. *Science and Public Policy*, 42(6), 789–810.
- Sweileh, W. M. (2020). Bibliometric analysis of scientific publications on “sustainable development goals” with emphasis on “good health and well-being” goal (2015–2019). *Globalization and Health*, 16(1), 1–13.
- Tahamtan, I., SafipourAfshar, A., & Ahamdzadeh, K. (2016). Factors affecting number of citations: A comprehensive review of the literature. *Scientometrics*, 107(3), 1195–1225.
- Thelwall, M. (2018). Do females create higher impact research? Scopus citations and Mendeley readers for articles from five countries. *Journal of Informetrics*, 12(4), 1031–1041.
- Thelwall, M. A. (2020a). Female citation impact superiority 1996–2018 in six out of seven english-speaking nations. *Journal of the Association for Information Science and Technology*, 71(8), 979–990.
- Thelwall, M. A. (2020b). Gender differences in citation impact for 27 fields and six English-speaking countries 1996–2014. *Quantitative Science Studies*, 1(2), 599–617.
- Thelwall, M. A., & Más-Bleda, A. (2020). A gender equality paradox in academic publishing: Countries with a higher proportion of female first-authored journal articles have larger first-author gender disparities between fields. *Quantitative Science Studies*, 1(3), 1260–1282.
- Thelwall, M. A., & Sud, P. (2020). Greater female first author citation advantages do not associate with reduced or reducing gender disparities in academia. *Quantitative Science Studies*, 1(2), 1283–1297.
- Thelwall, M., & Wilson, P. (2014). Regression for citation data: An evaluation of different methods. *Journal of Informetrics*, 8(4), 963–997.
- Tsay, My., & Li, Cn. (2017). Bibliometric analysis of the journal literature on women’s studies. *Scientometrics*, 113(2), 705–734.
- Tscharntke, T., Hochberg, M. E., Rand, T. A., Resh, V. H., & Krauss, J. (2007). Author sequence and credit for contributions in multiauthored publications. *PLoS Biology*, 5(1), e18.
- United Nations Development Programme (UNDP). (2020). *Human development report 2020*. Retrieved from <https://hdr.undp.org/content/human-development-report-2020>. Accessed 15 Sept 2021.
- van den Besselaar, P., & Sandström, U. (2016). Gender differences in research performance and its impact on careers: A longitudinal case study. *Scientometrics*, 106(1), 143–162.
- Wang, J., Thijs, B., & Glänzel, W. (2015). Interdisciplinarity and impact: Distinct effects of variety, balance, and disparity. *PLoS ONE*, 10(5), e0127298.
- Wang, X., Lv, T., & Hamerly, D. (2020). How do altmetric sources evaluate scientific collaboration? An empirical investigation for Chinese collaboration publications. *Library Hi Tech*, 38(3), 563–576.
- West, J. D., Jacquet, J., King, M. M., Correll, S. J., & Bergstrom, C. T. (2013). The role of gender in scholarly authorship. *PLoS ONE*, 8(7), e66212.
- World Economic Forum. (2019). *Global gender gap report 2020*. Retrieved from <https://www.weforum.org/reports/gender-gap-2020-report-100-years-pay-equality>. Accessed 15 Sept 2021.
- Wren, J., Kozak, K., Johnson, K., Deakyn, S., Schilling, L., & Dellavalle, R. (2007). The write position: A survey of perceived contributions to papers based on byline position and number of authors. *EMBO Reports*, 8(11), 988–991.
- Yegros-Yegros, A., Rafols, I., & D’este, P. (2015). Does interdisciplinary research lead to higher citation impact? The different effect of proximal and distal interdisciplinarity. *PLoS ONE*, 10(8), e0135095.
- Zeina, M., Balston, A., Banerjee, A., & Woolf, K. (2020). Gender and ethnic differences in publication of BMJ letters to the editor: An observational study using machine learning. *British Medical Journal Open*, 10(12), e037269.
- Zhang, L., Shang, Y., Huang, Y., & Sivertsen, G. (2021b). Gender differences among active reviewers: An investigation based on Publons. *Scientometrics*. <https://doi.org/10.31235/osf.io/4z6w8>
- Zhang, L., Sivertsen, G., Du, H., Huang, Y., & Glänzel, W. (2021a). Gender differences in the aims and impacts of research. *Scientometrics*, 126, 8861–8886. <https://doi.org/10.1007/s11192-021-04171-y>
- Zhang, L., Sun, B., Jiang, L., & Huang, Y. (2021c). On the relationship between interdisciplinarity and impact: Distinct effects on academic and broader impact. *Research Evaluation*, 30, 256.