https://doi.org/10.51408/issi2025_028

A Novel Bibliometric Algorithm Unveils the Prevalence and Significance of Gender Match in Research Mentorship Networks

David Campbell[1], Guillaume Roberge[2], Elisabeth Browning[3]

[1]*d.campbell@elsevier.com*, [2]*g.roberge@elsevier.com*, [3]*e.browning@elsevier.com*
Analytical Services, Elsevier (Canada)

Abstract

Gender homophily in research mentorship was investigated across a selection of countries and research fields over two decades by leveraging a novel tree algorithm that uncovers mentor–mentee links from genderized author profiles in Scopus. Despite a decrease in homophily for women relative to men, the overrepresentation of same-gender pairs remained much higher for women than for men in 2022. Only in fields where women were the dominant group was homophily in mentorship more pronounced for men than for women. Next, the contribution, relative to expectations, of same- and opposite-gender mentees to the tree index of their mentors was explored. This new metric quantifies the contribution of mentors to the future bibliometric performance of their mentees. Pairing with same-gender mentors was found to slightly and nearly systematically benefit the future bibliometric performance of women mentees across countries and research fields, regardless of their status as the underrepresented gender. While the positive impact on the performance of women mentees is small, the consistent pattern observed across countries and fields suggests that this is a genuine effect. The robustness of our findings across different contexts suggests that the availability of same-gender mentors is more critical for women than for men, due to women's lower representation in most areas of science. In contrast, the bibliometric performance of men mentees only appeared to benefit from a gender match in mentorship in the few subfields in which men are underrepresented. These results underscore the importance of gender match in research mentorship networks, particularly for women mentees, as well as critical aspects of the dynamics of research mentorship networks.

Introduction

Research mentorship refers to a multifaceted relationship where experienced researchers (mentors) guide less experienced individuals (mentees) in their personal and professional development. Such mentorship, when built on mutual respect and commitment, has been shown to significantly impact the professional trajectories of the next generation of scholars, enhancing their technical (e.g., research design, instrument use, data treatment and analysis) and non-technical skills (e.g., networking, authorship practices, fundraising, mentoring) in various aspects of research. This, in turn, contributes to the success of early career researchers (e.g., graduate students) as they move on to independent research careers (Guston, 1993; National Academies of Sciences-Engineering and Medicine, 2019).

Although research mentoring is often regarded as being within the remit of formal roles (e.g., supervisors), other researchers, such as collaborators, can fulfill informal mentorship roles through the provision of guidance and support (e.g., experience sharing, offering feedback) to less experienced individuals. Having multiple mentors can enrich a mentee's experience, notably by broadening perspectives and networking opportunities (Atenas et al., 2023; Gorela & Biloslavo, 2015).

A growing body of research suggests that same-gender mentor–mentee relationships within various academic settings are far more common than expected under a gender-neutral pairing assumption (Gallen & Wasserman, 2023; Moghe et al., 2021; Morales et al., 2018; Murphy et al., 2022; Schwartz et al., 2022). Furthermore, matching gender in mentor–mentee relationships, especially for women, could more effectively equip mentees for an academic career, yielding greater benefits than when mentored by someone of the opposite gender (Morales et al., 2018; Schwartz et al., 2022).

However, these studies were limited in geographic and/or disciplinary scope and were reliant on local surveys or databases (e.g., online mentoring platforms or databases indexing theses (COBISS)) to uncover and gather information on mentor–mentee relationships. Perhaps the most extensive analysis was performed by Schwartz and colleagues (2022) using data from the Academic Family Tree (https://www.academictree.org), though this was limited to the life sciences and had a predominantly US focus.

This study aims to confirm and assess the generalizability of existing evidence on the extent of gender homophily in research mentorship networks, as well as the potential benefits, in terms of bibliometric performance, of a gender match in mentorship. This is achieved by leveraging a novel tree algorithm and tree index (Roberge et al., 2024). By uncovering both formal and informal mentorship relationships from genderized author profiles in Scopus, the tree algorithm enables a large-scale examination of the dynamics of research mentorship networks over the past two decades, as well as across several countries and all fields of science to address the following questions:

- **Q1:** Are research mentorship networks gender homophilic?
- **Q2:** Is gender homophily in research mentorship networks more pronounced for the underrepresented gender?
- **Q3:** Is gender homophily among women researchers (usually the underrepresented gender) negatively correlated with their representation in research?
- **Q4:** Has gender homophily among women researchers (usually the underrepresented gender) declined as their representation in research increased over the past 20 years?

By capturing the average cumulative performance of a mentor's mentees as independent researchers later in their careers, the tree index enables assessing the benefits of same-gender mentorships on the subsequent bibliometric performance of mentees to address the following questions:

- **Q5:** Is gender match in mentorship beneficial to the bibliometric performance of mentees, as evidenced by the publications they produce independently of their mentors later in their career?
- **Q6:** Are performance gains from same-gender mentorships more pronounced for the underrepresented gender?

The paper concludes with a discussion of the implications of the study findings in light of the existing literature on gender homophily in research.

Methodology

Data source

The Scopus database produced by Elsevier includes abstracts and citation information from more than 90 million records covering all fields of science and technology, including the social sciences, arts and humanities. The September 1st 2023 snapshot of Scopus (Elsevier) was used to retrieve all necessary metadata on peer-reviewed scientific publications, mainly articles, conference papers, reviews, and short surveys published in book series, conference proceedings, or journals having valid ISSNs. Hereafter, these records are referred to as publications or papers.

Disambiguated author profiles

The novel tree algorithm and tree index used in this study rely on Scopus author IDs (AUIDs). The disambiguated AUIDs offer a clean view of a researcher's full publication history as indexed in Scopus. Using a large-scale "gold set" of over 10,000 manually cleaned author profiles, Campbell & Struck (2019) have shown that Scopus AUIDs enable robust conclusions in an evaluative context of groups of at least 500 researchers. They estimated the recall and precision of the AUIDs at, respectively, 98% and 96.9%.

Gender inference

The scale of the analysis pursued in this study, spanning several countries and scientific disciplines, required getting gender data for millions of Scopus authors. Collecting self-declared gender data at such a scale has never been undertaken and would not have been feasible within the time and budgetary constraints of the study. Therefore, the study team opted to infer binary gender, recognizing the limitation that this approach does not account for all gender groups. Additionally, had data been available for non-binary gender groups, including them could have risked identifying specific individuals as they represent a very small proportion of the population.
The Namsor API was used to infer the binary gender of all authors (covering mentors and mentees) in Scopus. Authors were classified as a man or a woman if the probability of being a man or a woman exceeded 85%. Pinheiro, Durning, & Campbell (2022) demonstrated that results were robust to changes in this gender assignation rule in a multivariate analysis of the relationship between gender and interdisciplinarity. Additionally, this is a well-established method that has been used in several rounds of *She Figures* by the European Commission (2021, 2024).

Tree algorithm

In this study, a tree algorithm was used to identify both formal and informal mentorship ties through an examination of key co-authorship patterns between senior and junior researchers, as summarized below.

The tree algorithm identifies mentor–mentee relationships from the publication history of Scopus AUIDs as follows (refer to Roberge et al. (2024) for more information):

- Senior researchers, defined as those with at least 10 years of publishing experience and at least one paper in the year under investigation, are identified.
- Researchers in Scopus are assigned as a potential mentor the senior researcher with whom they have co-published the most within their first five years as authors in Scopus. A minimum of 2 co-publications is required for mentor assignment, and ties can result in multiple mentors.
- Only mentees who later (after the first five years) published at least two papers independently from their potential mentor(s) are retained in the mentor's tree, ensuring it only includes mentees who had some success in a subsequent publishing career.

Roberge et al. (2024) utilized the E-Theses Online Service (EThOS)[1] of the National Library of the UK to validate the "mentoring" character of the identified mentor–mentee links using metadata from over 100,000 theses (mostly PhDs) awarded between 1980 and 2022. Excluding student–supervisor pairs from EThOS where the tree algorithm could not assign any mentor to the students matched to Scopus, the share of students linked to their correct supervisor(s) was 77% (ranging from 67% to 83% across disciplines, with lower accuracies observed in the Social Sciences and Humanities (SSH)). While validation using EThOS data showed that the tree algorithm frequently captures formal mentorship ties in the form of supervisor–student pairs, it is worth noting that it also captures other senior–junior interactions. In this paper, such collaborations between senior and junior researchers are assumed to have been accompanied by informal peer mentoring (e.g., experience sharing) known to occur in research collaboration networks. Future research could test this assumption by asking a sample of senior researchers to review the list of their non-student mentees as identified by the tree algorithm.
Using EThOS data, Roberge et al. (2024) also showed that in some disciplines, mainly in the SSH, supervisor–student pairs are not captured as frequently, likely due to coverage issues in Scopus (the bibliographic database used in this study). As a result, this study's findings for the SSH may not be as robust as for other disciplines.

Time series

Running the tree algorithm requires substantial computational resources. To analyze trends, data have therefore been generated for a limited set of years over the last 20 years, specifically in 2002, 2006, 2010, 2014, 2018, and 2022. In any of these years, the tree algorithm accounted for the full publication history of a mentor, and its mentees, up to that year (inclusive).

[1] https://bl.iro.bl.uk/concern/datasets/308c54ce-31b1-4cb1-b257-7b288a3c7926?locale=en

Country and field coverage

For each of the above years, the mentors have been limited to authors who qualified as senior and actively published at least one paper in the corresponding year. Mentors were uniquely assigned to the country and subfield in which they published most of their publications up to that year.[2] Although researchers might have visited more than one country, the homophily signature linked to a given mentor should mostly reflect the situation in the main country of affiliation. When aggregating across a country's mentors, the extent to which they contributed to homophily is expected to converge to the country's main pattern minimizing noise from the secondary countries of mentors.

To test the robustness of the study findings, as well as to assess their generalizability, the analyses were repeated for 38 countries (EU27 members plus Argentina, Australia, Brazil, Canada, Egypt, India, Japan, Mexico, South Africa, United Kingdom, and the United States)[3] overall in Scopus and by main field of research based on Science-Metrix's classification.[4]

Some analyses were repeated for a small group of subfields with a majority of woman mentors (i.e., Gender Studies, Nursing, Nutrition & Dietetics, Developmental & Child Psychology, Public Health and Social Work). This was done to assess whether gender differences in mentorship homophily and associated performance gains differ from the dominant patterns where male researchers are in the majority.

Homophily indicator

To test hypotheses 1 to 4, the study examined, across the selected fields, countries, and years, the extent to which the frequency of same-gender links departs from expectation under a gender-neutral (random) pairing assumption by gender of the mentors.

For each mentor in a given year, all prior mentees up to that year are considered regardless of originating country(ies), as students may come from abroad. Therefore, the expected share of women and men mentees in a given field, country and year for each mentor, regardless of gender, is based on the pooled set of mentees of a country's mentors, including those from abroad, in the corresponding field and year. As an example, if 44% of the pooled mentees of women and men mentors in Canada are women, one would expect 44% of the mentees of women mentors to be women if the assortment of mentors and mentees was gender neutral. A share of woman–woman links above 44% would denote a homophilic network from the perspective of women mentors. In Canada, 60% of the mentees of women mentors are women leading to a positive deviation of 36% relative to expectation (homophily for women mentors = (0.60/0.44) - 1 = 0.36). The same approach was applied in

[2] In the rare case of ties, mentors were randomly assigned to one of the tied countries and were assigned to all the tied subfields. We will further test the impact on the study conclusions by either excluding mentors or assigning them to multiple countries in the case of ties.

[3] China was not included due to issues in assigning gender to authors.

[4] https://www.science-metrix.com/classification/

exploring the extent of deviation from expectation for man–man links (homophily for men mentors = (0.61/0.56) - 1 = 0.09).

Tree index

To address questions 5 and 6, the contribution of same- and opposite-gender mentees to the tree index of their mentors was investigated across the selected fields, countries, and years.

The tree index is a new subfield- and year-normalized metric designed to quantify the impact of mentors on the bibliometric performance of the next generation of scientists. For each mentor, this composite indicator accounts for the cumulative volume and impact of his or her mentees' publications, as well as the size of their co-authorship network, as they go on to independent careers (refer to Roberge et al. (2024) for more information). As such, each mentee contributes a certain share to the mentor's tree index.

If mentees of women mentors consist, on average, of 61.21% women in Canada, and the average share of these mentors' tree index that is attributable to women mentees equals 60.76%, the tree index departure from expectation for woman–woman (mentee–mentor) pairings would equal -0.7% (0.608/0.612 - 1). Applying the same approach to other pairing types produces the following deviations for Canada in Scopus (2022): -0.7% woman–woman, -2.8% woman–man, 1.7% man–man, and 1.1% man–woman.

As observed for Canada, woman mentees are likely to contribute less than expected to their mentors' tree index regardless of the gender of their mentors because of gender inequalities in research (European Commission, 2021). The opposite applies to men mentees. Nevertheless, by comparing the average departure from expectation in the contribution of mentees to the tree index of their women and men mentors, the study enables an assessment of whether same-gender mentorships are associated with performance gains for both women and men mentees.

Note that expectation assumes all else is equal even if not the case (e.g., publishing age of a mentor's mentees). This is later accounted for in interpreting the results.

Results

Descriptive statistics on the study data set

Among the selected countries (all fields combined), the share of mentors and mentees with unknown gender was small, with no major implication for the study findings (Figure 1, left). As expected, the share of women mentors and mentees has been steadily increasing over the past two decades (Figure 1, right). The share of women among the approximately 2.4 million mentees was unsurprisingly higher than their share among the approximately 0.6 million mentors in 2022. This is due to the well-known leaky pipeline and glass ceiling phenomena in academia, whereby a smaller share of women researchers reaches senior levels (European Commission, 2021).

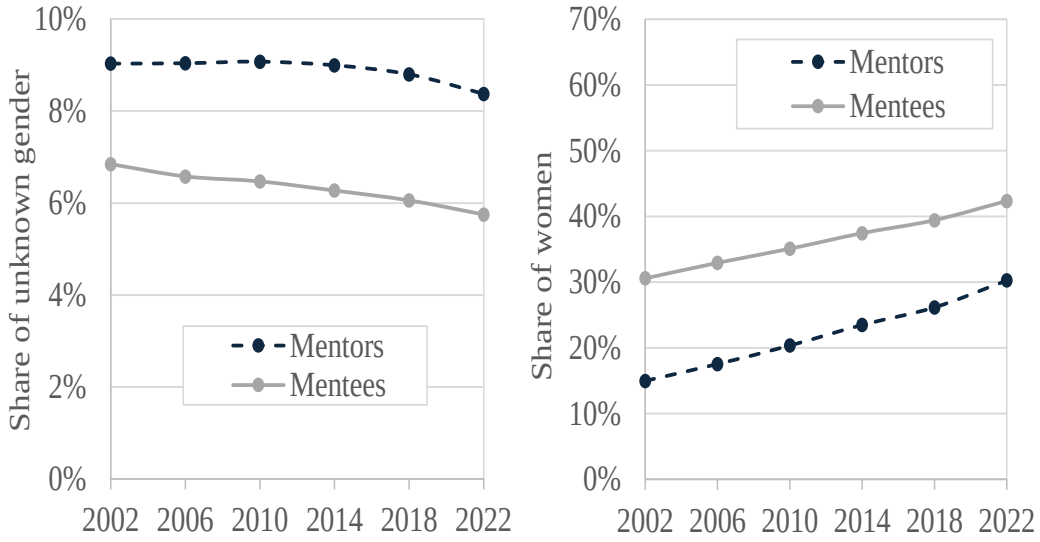

Figure 1. Trends in the average share of unknown gender (left) and women (right) among mentors and mentees of the selected countries, 2002–2022.

Note: The share of women is calculated excluding unknowns.

Questions 1 to 4

The percent deviation in the share of same-gender mentees relative to neutral expectation (random assignment without regard to gender) is depicted by country and gender of the mentors in Figure 2, for all fields combined in 2022. Some key patterns emerge. First, over the course of mentors' careers, the share of same-gender mentees is systematically higher than expected for both men and women mentors active in 2022. Second, and more interestingly, the degree of homophily is systematically more pronounced for women, usually the underrepresented gender, than for men mentors. On average across the selected countries, woman−woman links are overrepresented by 43% relative to expectation, versus 11% for man−man links. There is also a moderate negative correlation (-0.53) between the share of women mentees and the homophily of women mentors across countries (data not shown). Thus, with greater representation of women in research, one might expect a decrease in homophily within research mentorship networks, a pattern that is indeed observed in the study's results.

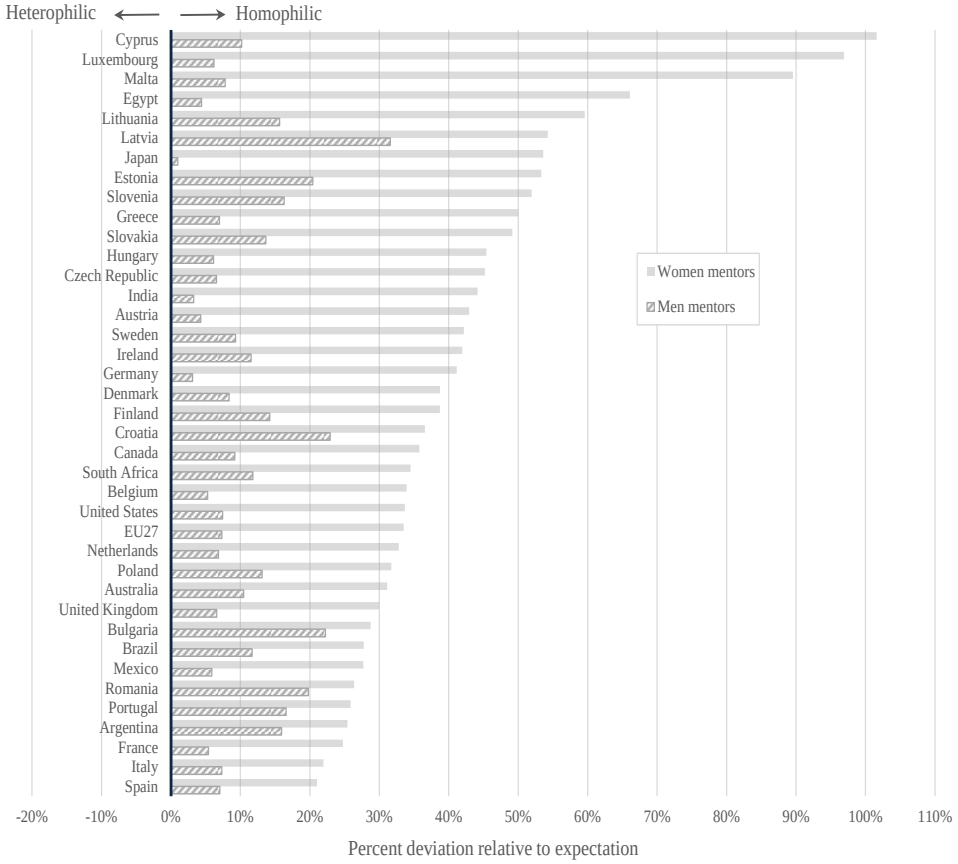

Figure 2. Percent deviation in the share of same-gender mentees relative to expectation in the overall research mentorship network, by country and gender of the mentor (2022).

Across the selected countries (all fields combined), the extent of homophily systematically decreased for women (average CAGR of -4.0%) as their representation increased over the past two decades (Table 1). This pattern was matched by an opposite trend of the same magnitude for men (average CAGR of +3.8%). The increase for men was also systematic across countries.

Table 1. Trends in percent deviation in the share of same-gender mentees relative to expectation in the overall research mentorship network, by country and gender of the mentor, 2002–2022.

	Women mentors				Men mentors			
Country	**2002**	**2022¥**	**CAGR**	**Trend***	**2002**	**2022**	**CAGR**	**Trend***
Egypt	185%	66%	-5.0%		3.1%	4.4%	1.7%	
Japan	59%	54%	-0.5%		0.4%	1.0%	5.0%	
Slovenia	121%	52%	-4.1%		8.1%	16.3%	3.6%	
Greece	157%	50%	-5.6%		3.9%	7.0%	2.9%	
Slovakia	113%	49%	-4.1%		8.3%	13.7%	2.5%	
Hungary	93%	45%	-3.5%		3.6%	6.2%	2.6%	
Czech Republic	107%	45%	-4.2%		5.8%	6.6%	0.7%	
India	80%	44%	-2.9%		2.3%	3.3%	1.7%	
Austria	123%	43%	-5.1%		1.2%	4.3%	6.6%	
Sweden	84%	42%	-3.4%		3.0%	9.3%	5.9%	
Ireland	138%	42%	-5.8%		3.3%	11.6%	6.5%	
Germany	99%	41%	-4.3%		0.8%	3.1%	7.2%	
Denmark	77%	39%	-3.4%		2.9%	8.4%	5.5%	
Finland	68%	39%	-2.8%		7.1%	14.2%	3.6%	
Croatia	82%	37%	-3.9%		13.3%	22.9%	2.8%	
Canada	86%	36%	-4.3%		3.2%	9.2%	5.4%	
South Africa	142%	34%	-6.8%		4.0%	11.8%	5.5%	
Belgium	87%	34%	-4.6%		2.1%	5.3%	4.8%	
United States	66%	34%	-3.3%		2.8%	7.4%	5.0%	
EU27	71%	33%	-3.7%		3.5%	7.3%	3.7%	
Netherlands	68%	33%	-3.6%		1.1%	6.9%	9.6%	
Poland	55%	32%	-2.7%		9.9%	13.1%	1.4%	
Australia	75%	31%	-4.3%		3.6%	10.5%	5.4%	
United Kingdom	62%	30%	-3.6%		2.3%	6.6%	5.4%	
Bulgaria	57%	29%	-3.4%		13.8%	22.3%	2.4%	
Brazil	70%	28%	-4.5%		7.8%	11.7%	2.0%	
Mexico	94%	28%	-5.9%		5.2%	5.9%	0.7%	
Romania	63%	26%	-4.2%		10.5%	19.8%	3.2%	
Portugal	73%	26%	-5.0%		10.6%	16.6%	2.3%	
Argentina	57%	25%	-4.0%		14.5%	15.9%	0.5%	
France	42%	25%	-2.6%		3.4%	5.4%	2.3%	
Italy	36%	22%	-2.4%		4.1%	7.3%	2.9%	
Spain	45%	21%	-3.7%		3.9%	7.0%	3.0%	

Note: ¥Countries are sorted based on the extent of homophily for women mentors. *The trends are on a common scale to show that the absolute magnitude of change in homophily is smaller for men than women mentors despite the average magnitude of their CAGR being very similar, yet in opposite direction. Some countries are excluded due to too few observations.

Finally, when disaggregating the data presented in Figure 2 by field of science (data not shown), 99% of observations confirm the tendency towards same-gender pairings for both women and men mentors. The pattern is also more pronounced for women than men mentors 86% of the time. Interestingly, among the few subfields in which most mentors are women (i.e., Gender Studies, Nursing, Nutrition & Dietetics, Developmental & Child Psychology, Public Health and Social Work) and for which there is enough data (selected countries pooled) to analyse gender homophily in mentorship, the pattern of greater same-gender pairing was inverted. In these subfields (grouped), in 2022, a tendency to same-gender pairing is more pronounced for men (+25% deviation) than for women (+7%) (Figure 3).

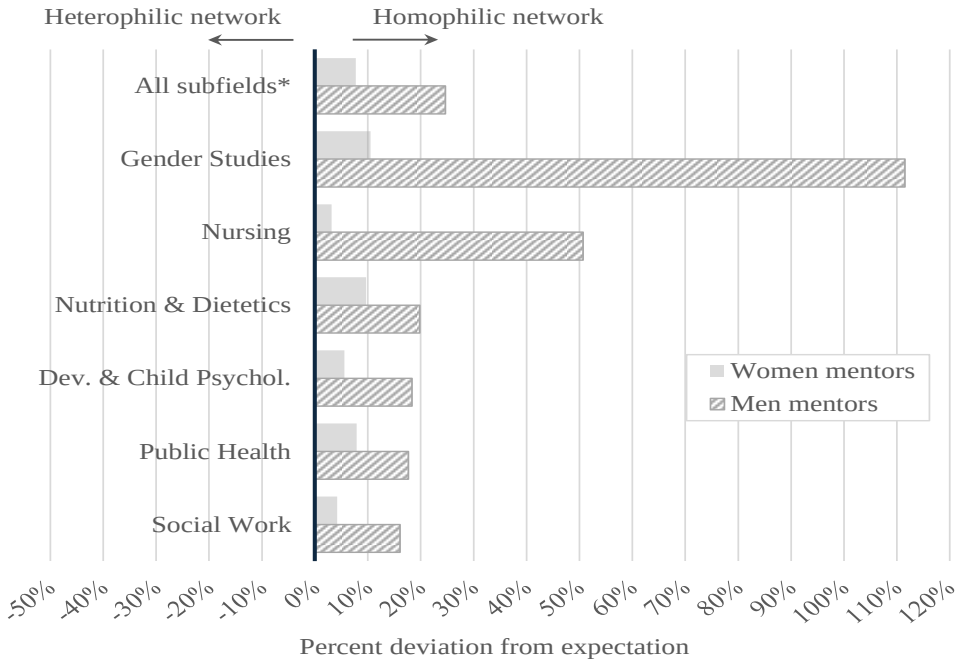

Figure 3. Percent deviation from expectation in the share of same-gender mentees in the overall research mentorship network, by subfield in which most mentors are women and gender of the mentor (2022).

Note: *All subfields are all those with a majority of women mentors grouped.

Questions 5 & 6

Accounting for those mentors who were actively publishing in each of the selected countries in 2022 (all fields combined), Figure 4 presents the extent to which the contribution of women and men mentees to the tree index of women and men mentors deviated from expectations. The key gender differences highlighted below and in Figure 4 do not appear to be due to gender differences in the average age of a mentor's mentees (data not shown):

- As anticipated, due to long-standing inequalities in research, women mentees contributed less than expected to the tree index of their mentors regardless of their mentors' gender. The opposite was observed for men mentees. Although

these departures from expectation are of a small magnitude, they are nearly systematic with only a handful of exceptions (i.e., Latvia, Estonia and Luxembourg).

- The underperformance of women mentees relative to expectation when paired with mentors of the same gender is, on average, roughly less than half that observed when they are paired with mentors of the opposite gender. This pattern is also systematic across all selected countries in all fields combined (except for Luxembourg and Lithuania) and was repeated in 75% of all country−field combinations (data not shown).
- A similar result is not consistently observed for men mentees for whom being paired with a mentor of the same gender only equated to performance gains for about half of the countries examined (18 out of 38 countries).

Despite the small magnitude of observed deviations from expectations, their consistency across countries and fields in 2022 suggests that same-gender mentors for junior women researchers offer some benefits, even if only slightly. This consistency warrants a deeper investigation into the root causes of this finding to uncover ways in which men mentors could better support women mentees.

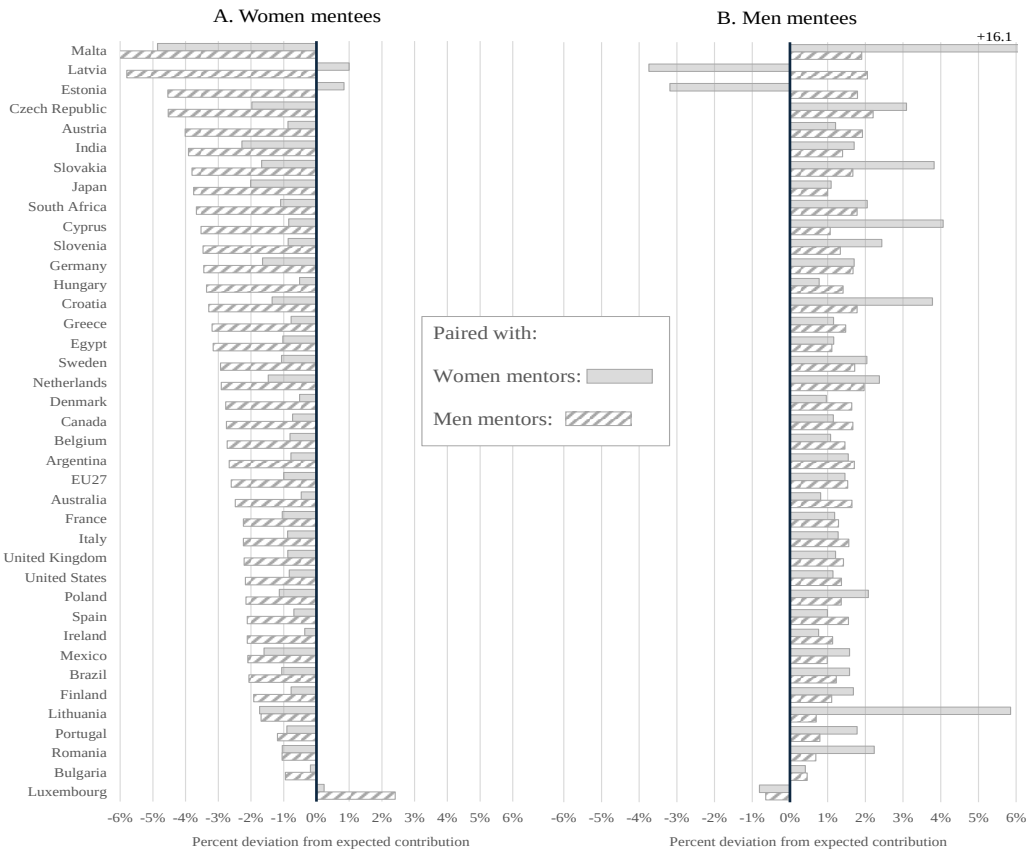

Figure 4. Percent deviation from expected contribution of women/men mentees to the tree index of women/men mentors in the overall research mentorship network, by country (2022).

Although the main patterns in network homophily were inverted for subfields with a majority of women mentors (see Figure 3), the patterns were not fully inverted in terms of the contribution of mentees to the tree index of mentors (Figure 5). The closest match to perfect inversion was observed in Gender Studies, where men mentees underperformed and women mentees overperformed regardless of the gender of their mentors. In that subfield, the negative deviation for men mentees was also less pronounced when paired with men than women mentors. However, in all other subfields, women mentees still underperformed relative to expectation and benefited from being paired with women mentors, although the effect sizes were still very small. Interestingly, in these subfields, even though men mentees systematically overperformed relative to expectation, they also appeared to systematically benefit from being paired with same-gender mentors, a pattern that was not observed in areas with a majority of men mentors.

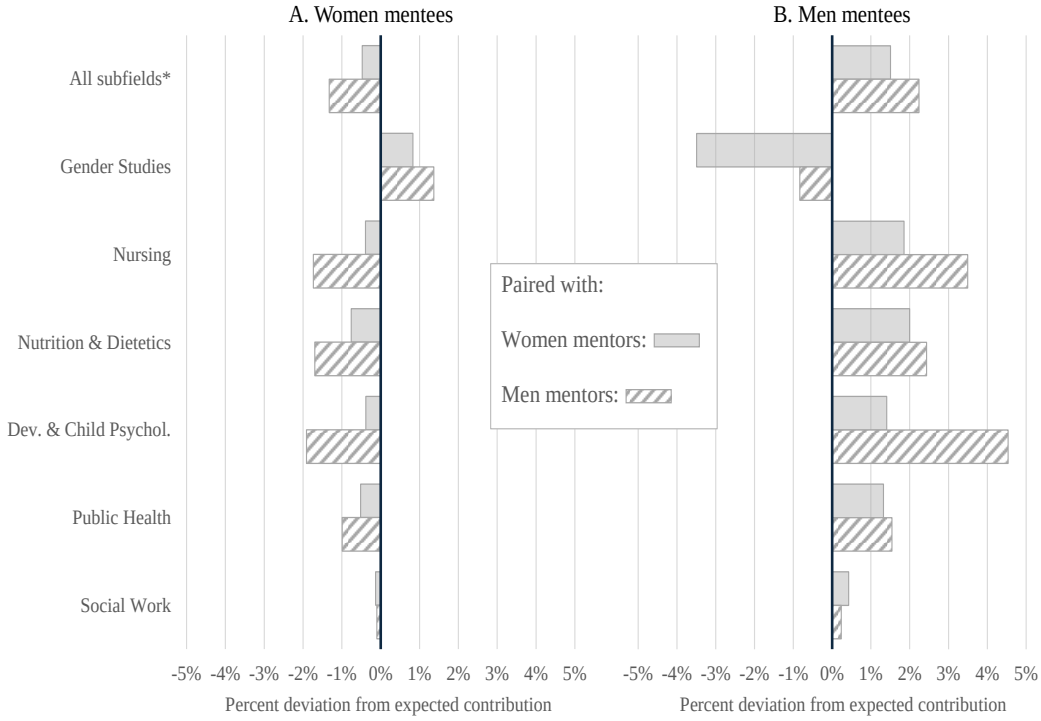

Figure 5. Percent deviation from expected contribution of women/men mentees to the tree index of women/men mentors in the overall research mentorship network, by subfields with a majority of women mentors (2022).

Note: *All subfields are all those with a majority of women mentors.

Discussion

Research mentorship networks were shown to be gender homophilic across nearly all country–field combinations for both women and men, generalizing findings from smaller-scale studies (Gallen & Wasserman, 2023; Moghe et al., 2021; Morales et al., 2018; Murphy et al., 2022; Schwartz et al., 2022). This suggests that mentors,

regardless of their gender, likely play an important role in attracting same-gender mentees to research and mentoring them later on.

Results also confirm that gender homophily in mentorship is usually more pronounced for the underrepresented gender. Homophily was more pronounced among women, typically the underrepresented gender, than among men researchers in 86% of the country–field combinations examined. This pattern was also observed in research collaboration networks (Hajibabaei et al., 2022; Wang et al., 2023), although not systematically (Kwiek & Roszka, 2021). Additionally, this pattern was inverted (i.e., greater homophily among men than women) in the few subfields where most mentors are women and for which there was enough data to analyse gender homophily in mentorship.

The extent of gender homophily among women was also negatively correlated with their representation in research. Women were more likely to pair with other women in country–field combinations where they were more underrepresented. Although the correlation was moderate, additional observations suggest that the more pronounced homophily among women could be driven by their status as the underrepresented group. Over the past two decades, the extent of homophily among women decreased with their increased representation in research, and the extent of homophily was found to be greater among men in the few areas of science where they are underrepresented (i.e., Gender Studies, Nursing, Nutrition & Dietetics, Developmental & Child Psychology, Public Health and Social Work).

The study's results therefore suggests that the availability of same-gender mentors is more critical for women than for men, due to women's lower representation in most areas of science. The importance of gender match in mentorship is further exemplified by the nearly systematic, albeit small, positive impact of same gender mentors on the later research performance of women mentees. Although the bibliometric performance of women mentees appeared to benefit from same-gender mentors regardless of their status as the underrepresented gender, men mentees only appeared to benefit from same-gender mentors in the few areas in which they were underrepresented. This is consistent with prior studies reporting benefits, especially for women, of matching gender in mentor−mentee relationships (Moghe et al., 2021; Morales et al., 2018; Schwartz et al., 2022).

It seems possible that members of underrepresented groups might seek out others with similar experiences and concerns about potential issues with the majority group. However, as diversity increases, these obstacles might diminish. Research by Bai, Ramos, & Fiske (2020) has shown that "as diversity increases, people paradoxically perceive social groups as more similar," possibly leading to fewer stereotypes.

In turn, as gender becomes more equally represented in science, the extent of homophily would be expected to decrease. This underscores the importance of retaining senior women researchers, not only for their direct contributions to science but also for their indirect mentorship contributions via their mentees.

Given the study's limitations as detailed throughout the methods section (e.g., use of inferred binary instead of self-declared non-binary gender, need to further validate the "mentoring" connection of uncovered informal peer mentoring ties), further work relying on a mix of qualitative and quantitative approaches would be warranted to

confirm our findings on homophily and its implications for mentees. However, the robustness of our findings is supported by the parallel results obtained across countries and fields, reinforcing the validity of our conclusions.

These findings should guide policymakers in initiatives aimed at encouraging the greater participation of women in science. For instance, interventions to increase the participation of women in science, especially in countries or fields, such as science, technology, engineering, or math where they are heavily underrepresented, appear highly relevant considering the study's results. Interventions directed towards increasing the retention rates of women as they advance through academic careers may be particularly effective. In these cases, concomitant interventions to strengthen the mentoring skills of men towards women may also be warranted, and further research could help uncover the main levers for intervention.

Altogether, increasing gender diversity in research, and in research teams, should be the ultimate target as several recent studies underscore the unique value of mixed-gender teams in fostering novel, disruptive, and influential scientific discoveries (Hajibabaei et al., 2022; Yang et al., 2022; Zhang et al., 2024).

Acknowledgements

The authors would like to thank senior management at Elsevier, in particular Nicolien van der Linden, M'hamed el Aisati and Olivier Dumon, for their interest and support in leveraging the novel tree algorithm and tree index to address key questions of high relevance to promoting an inclusive and diverse research culture at all stages of an academic career. This research was funded through mandated work conducted for the European Commission Directorate General for Research and Innovation.

Author contributions

Conceptualization: David Campbell
Methodology: David Campbell, Guillaume Roberge
Data curation: David Campbell, Guillaume Roberge
Formal Analysis: David Campbell
Writing – original draft: David Campbell
Writing – review & editing: Elisabeth Browning, David Campbell, Guillaume Roberge

Competing interests

The authors declare no competing interests.

Declaration of generative AI and AI-assisted technologies in the working process

The authors used Scopus AI to assist the literature search. During preparation of the manuscript, the authors used Microsoft Copilot, built upon OpenAI's GPT-4, to improve readability and language of certain sentences. The authors carefully reviewed and edited the proposed changes as needed and take full responsibility for the content of the publication.

References

Atenas, J., Nerantzi, C., & Bussu, A. (2023). A Conceptual Approach to Transform and Enhance Academic Mentorship: Through Open Educational Practices. *Open Praxis*, *15*(4), 271–287. https://doi.org/10.55982/OPENPRAXIS.15.4.595

Bai, X., Ramos, M. R., & Fiske, S. T. (2020). As diversity increases, people paradoxically perceive social groups as more similar. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(23), 12741–12749. https://doi.org/10.1073/PNAS.2000333117/SUPPL_FILE/PNAS.2000333117.SAPP.PDF

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

European Commission. (2021). *She Figures 2021*. https://research-and-innovation.ec.europa.eu/knowledge-publications-tools-and-data/publications/all-publications/she-figures-2021_en

European Commission. (2024). *She Figures 2024*. Upcoming

Gallen, Y., & Wasserman, M. (2023). Does information affect homophily? *Journal of Public Economics*, *222*, 104876. https://doi.org/10.1016/J.JPUBECO.2023.104876

Gorela, K., & Biloslavo, R. (2015). Relationship between Senior and Junior Researcher: Challenges and Opportunities for Knowledge Creating and Sharing. In P. Diviacco, P. Fox, C. Pshenichmy, & A. Leadbetter (Eds.), *Collaborative Knowledge in Scientific Research Networks* (pp. 90–125). IGI Global. https://doi.org/10.4018/978-1-4666-6567-5.CH006

Guston, D. H. (1993). Mentorship and the Research Training Experience. In National Academy of Sciences, National Academy of Engineering and Institute of Medicine (Ed.), *Responsible Science: Ensuring the Integrity of the Research Process: Volume II*. (pp. 50–65). National Academies Press (US). https://www.ncbi.nlm.nih.gov/books/NBK236193/

Hajibabaei, A., Schiffauerova, A., & Ebadi, A. (2022). Gender-specific patterns in the artificial intelligence scientific ecosystem. *Journal of Informetrics*, *16*(2), 101275. https://doi.org/10.1016/j.joi.2022.101275

Kwiek, M., & Roszka, W. (2021). Gender-based homophily in research: A large-scale study of man-woman collaboration. *Journal of Informetrics*, *15*(3), 101171. https://doi.org/10.1016/J.JOI.2021.101171

Moghe, S., Baumgart, K., Shaffer, J. J., & Carlson, K. A. (2021). Female mentors positively contribute to undergraduate STEM research experiences. *Volume 16, Issue 12 December, 16*(12 December). https://doi.org/10.1371/journal.pone.0260646

Morales, D. X., Grineski, S. E., & Collins, T. W. (2018). Effects of gender concordance in mentoring relationships on summer research experience outcomes for undergraduate students. *Volume 102, Issue 5, Pages 1029 - 1050*, *102*(5), 1029–1050. https://doi.org/10.1002/sce.21455

Murphy, M., Record, H., Callander, J. K., Dohan, D., & Grandis, J. R. (2022). Mentoring Relationships and Gender Inequities in Academic Medicine: Findings from a Multi-Institutional Qualitative Study. *Academic Medicine*, *97*(1), 136–142. https://doi.org/10.1097/ACM.0000000000004388

National Academies of Sciences-Engineering and Medicine; Policy and Global Affairs; Board on Higher Education and Workforce; Committee on Effective Mentoring in STEM. (2019). The Science of Mentoring Relationships: What Is Mentorship? In M. L. Dahlberg & A. Byars-Winston (Eds.), *The Science of Effective Mentorship in STEMM*

(pp. 33–50). National Academies Press (US). https://doi.org/10.17226/25568

Pinheiro, H., Durning, M., & Campbell, D. (2022). Do women undertake interdisciplinary research more than men, and do self-citations bias observed differences? *Quantitative Science Studies*, *3*(2), 363–392. https://doi.org/10.1162/qss_a_00191

Roberge, G., Campbell, D., Browning, E., Dong, D., Khayat, P., El Aisati, M., & Dumon, O. (2024). Tree Index: a new widescale indicator on contribution to mentorship. *Proceedings of the 2024 STI Conference, Berlin, 18-20 September*. https://zenodo.org/records/13987541

Schwartz, L. P., Liénard, J. F., & David, S. V. (2022). Impact of gender on the formation and outcome of formal mentoring relationships in the life sciences. *Volume 20, Issue 9, 20*(9). https://doi.org/10.1371/journal.pbio.3001771

Wang, Y. S., Lee, C. J., West, J. D., Bergstrom, C. T., & Erosheva, E. A. (2023). Gender-based homophily in collaborations across a heterogeneous scholarly landscape. *PLoS ONE*, *18*(4). https://doi.org/10.1371/journal.pone.0283106

Yang, Y., Tian, T. Y., Woodruff, T. K., Jones, B. F., & Uzzi, B. (2022). Gender-diverse teams produce more novel and higher-impact scientific ideas. *Proceedings of the National Academy of Sciences of the United States of America*, 119(36), e2200841119. https://doi.org/10.1073/PNAS.2200841119/SUPPL_FILE/PNAS.2200841119.SAPP.PDF

Zhang, M. Z., Wang, T. R., Lyu, P. H., Chen, Q. M., Li, Z. X., & Ngai, E. W. T. (2024). Impact of gender composition of academic teams on disruptive output. *Journal of Informetrics*, 18(2), 101520. https://doi.org/10.1016/j.joi.2024.101520