Age, Gender, and International Author Networks in the Earth and Space Sciences: Implications for Addressing Implicit Bias

Brooks Hanson[1], Paige Wooden[1], Jory Lerback[2]

1. American Geophysical Union, Washington, DC 20005
2. University of Utah, Salt Lake City, UT
Correspondence to bhanson@agu.org

Abstract
Author networks play a key role in doing science. Developing networks is critical for career advancement in a wide variety of ways, and differences in networks may be a core reason for persistence of implicit gender bias. Combining the AGU Fall Meeting abstracts from 2014-2018 with self-identified AGU member data on birth year and gender provides a large database of more than 400,000 unique co-author interactions that we use to examine author networks by age, gender, and country. Age data are necessary to disambiguate the effect that a historic lack of women in the Earth and space science. The data show that women's networks are closer to those expected from the age-gender distribution of the overall membership; whereas networks of men include more men than expected, although women are also interacting with men of similar age more than expected from the membership. Women's networks are also less international than their male colleagues in most age cohorts. These differences start in the youngest age cohort. These data indicate that addressing implicit bias requires efforts at encouraging and developing more balanced author networks, particularly in early-career scientists. Recent work suggest that this will also improve science outputs.

1. Introduction

Science is increasingly collaborative, and the best and most impactful science is regularly done in teams (Hall et al., 2018, provide a recent review; Wu *et al.*, 2019). As such, individual success and career advancement increasingly depend on developing and fostering a broad network of collaborators. Large collaborator networks have a myriad of additional benefits, including in recommending students and postdocs, help with research, providing international introductions, and more (e.g., Teplitskiy *et al.*, 2018).

At the same time implicit bias is increasingly being recognized in the practice of science—mostly around gender (often inferred from binary sex data) as that is the easiest to garner sufficient data for significance—but likely extending to age, ethnicity, and other forms of identities. This bias is manifested in lower invitations to serve as reviewers, acceptance rate in some disciplines (but not all), invitations to speak at conferences, hiring and promotion, award nominations, and many more activities (West et al., 2013; Holmes, 2015; Lerback and Hanson, 2017; Helmer *et al.*, 2017; Ford *et al.*, 2018; Roper, 2019; Fox and Paine, 2019).

Several studies have shown that including women in key roles, such as editors or session conveners, reduces gender bias (Helmer, 2017; Lerback and Hanson, 2017). Reminders of bias and training also help, but these measures do not always fully correct the observed differences, and recent studies have shown that biases persist despite these efforts. For example, several recent studies have shown a small (ca. 5%) but significant difference in the nominations of women by men for roles—for example for reviewers by a male editor or speakers by a male session convener (Ford *et al.*, 2017; Hanson and Lerback, 2017) even after these interventions.

One reason that addressing bias may be difficult and training alone not sufficient is that the bias may be explicitly manifested in the networks that scientists form starting early in their careers, which creates a persistent long-lasting effect. Suggestions and recommendations, for example, for reviewers or speakers and especially award nominations, often start from close networks, or follow along them, and bias would then reflect differences in the makeup of networks among men and women. If so, addressing the bias would require fostering diverse networks especially early in careers when close and lifelong collaborations form and continuing to do so over that career, a significantly more difficult and complex challenge than providing implicit bias training (which is still important).

To examine whether and how networks of men and women were different, and thus might be contributing to persistent gender bias, we analyzed co-authors by gender and age using a large database of abstracts for the AGU Fall Meeting for the past 5 years (2014-2018). These datasets, merged with the AGU member database, provide self-identified information on age, gender, and country of residence of authors. In general, such information, particularly around age, has not been available in most prior studies, many of which also use name-gender algorithms where the uncertainty or error rate may be close to the size of signals of bias. Age data are necessary to disambiguate the effect of historical underrepresentation of women and other demographic groups in the sciences. Such data also allow us to assess how co-author networks might vary or develop over career stages or with time. We also examine the geographical composition of co-author networks and how these vary by age and gender. While professional networks are certainly broader than co-authors alone, these are particularly strong and trusted connections that persist through careers and are often used for recommendations, career references, and more. Broader geographical connections can help science in other ways too, for example, through science diplomacy, capacity building, and improved data access.

2. Approach

The AGU Fall Meeting is one of the largest annual scientific conferences worldwide. Recent meetings have had close to or more than 25,000 abstracts representing ~100,000 distinct authors each year from ~130 countries. Membership in AGU is generally required to submit an abstract, and most members have self-declared their birth year and gender. Additional data for some non-members including abstract co-authors are included in the AGU databases. Random checks indicate that these data are highly accurate (Lerback and Hanson, 2017). From 2014 to 2018, a 5-year period, these provide a data set with 417,632 unique co-author pairs with age, gender, and additional demographic information (or 835,264 bi-directional connections), representing 70,519 unique authors where we were able to match age and gender. The total number of author pairs on abstracts over this time period is 1,239,473, so we were able to match about 67% with age and gender. The AGU Fall Meeting includes some abstracts from high-school students and early undergraduates (<20 years old) as well as scientists who were greater than 90 years old; we excluded these from the analysis (both co-author connections from them and with them) as these groups are too small for statistical power. In all cases, age was adjusted for the time of the first activity; thus, we assigned author connections that persisted over several or all years to their youngest respective ages in this dataset. We grouped ages by decades; thus, we counted researchers whose ages and author networks spanned two decadal cohorts over the 5-year dataset in the younger cohort. Although we focused on unique pairs over the 5 years of meeting abstracts, analysis of all pairs across all 5 years or pairs in individual years separately yielded similar results (with higher or lower sample sizes and significance) and indicates that considering the age of first interaction over this multi-year sample is representative.

Most AGU members and abstract authors were from the United States, Japan, and Europe; participation from China has been growing (from ~2000 abstract authors in 2014 to ~5000 in 2018). Thus our analysis is strictly for AGU members or other participants in society activities who have provided age and gender data, and we have likely undercounted especially some international collaborations with or among non-AGU members, as abstract authors who are not members are less likely to be from the US, Japan, and Europe. As the historical participation of women has been lower in many countries outside of North America and Europe, we expect that the observed gender differences are thus minima. Here we focus on the unique author networks rather than team size, persistence, or other dynamics. A related paper examines the relation between these networks and acceptance rates and citations in AGU Journals (Lerback et al., submitted).

3. Results

3.1 Gender-Age Networks

We analyzed author networks by gender and decadal age cohorts. Figure 1 shows the cumulative author networks by age and gender for each age-gender cohort (Fig. 1; data are in Table 1 in Hanson *et al.*, 2019). The networks are broadly similar to each other and reflect the overall AGU membership population distribution with slight but important differences. The proportion of women in the AGU member population is close to that of the employed geoscientists in the U.S. All age-gender cohorts show the most co-authors with mid-career (30-59 year old) scientists, as these represent some of the most active researchers involved in both mentoring and forming larger research teams and are the largest AGU member cohort and represent the largest number of scientists presenting at the Fall Meeting (Fig. 1). Because of historical bias against women, there are relatively fewer older women in the Earth and space sciences, and thus fewer senior women co-authors of other authors of any age and gender. As expected, younger co-authors of authors in all age groups include more women. About 10,000 of the ca. 417,000 unique co-author connections were between two researchers in their 20s (or about 2.4%), whereas about 40,000 of the unique connections were between researchers in their 40s, or about 10%.

Although the patterns are broadly similar, some differences in networks by age are evident that also help understand gender and international differences. In particular, men and women across all age cohorts tend to have had more co-authors--both men and women--of their same age (the bars in Fig. 1 show the overall age-gender distribution of the full database). This is particularly the case in the older cohorts (50s and higher). The youngest cohorts, both men and women, have proportionally fewer co-author relations with the older cohorts, and vice versa, than expected based on the overall member population.

Although all networks for age-gender cohorts are broadly similar, the data show that women's networks for any age-cohort have relatively—typically several percent—more women as co-authors than their male counterparts (Fig. 2) and that this difference persists across most age cohorts. This includes the early-formed networks between researchers in their 20s and 30s. Interestingly, the co-author network for women in any age cohort is close to the overall AGU membership distribution; in contrast, for men, women are underrepresented by about 5% in each age cohort. This difference is highly significant for most networks (see Table 1 in Hanson *et al.*, 2019), except those involving scientists in their 70s and 80s, where there are few women in particular. The difference is also close to that seen above in recommendations for reviewers and invited speakers by male and female leaders (Ford *et al.*, 2018).

We can use the number of abstracts per author and average number of co-authors by age as a proxy for the development of networks by researchers over time (Fig. 3 and Table 2 in Hanson *et al.*, 2019). The data show that the number of abstracts essentially doubles by decade for authors between their 20's to 40's, as does the number of co-authors per person. This increase is similar for both men and women. Given the 5-year time window of our analysis, this would not reflect the full co-author network over a career, which would be expected to be larger, but does give a sense of how active networks might develop. Overall men were an author on 5.9 abstracts in the 5-year period whereas women were on 4.2 abstracts on average. This difference partly reflects that the average age of men as authors is older than the average age for women, but even within each cohort, men are authors on an average of 1.7 more abstracts over this time period (Fig. 3).

3.2 International teams

We analyzed authors' international networks using the country in author addresses on their earliest abstract of our dataset. In this analysis, we considered the 15 top countries individually and then grouped other countries based on region (Africa, rest of Asia, Europe, and South America). The data below for authors in each age cohort are for their co-authors across all age groups including their own. Of the 149,125 distinct abstract authors, 95% had a known country (144,708). Of those, 47% were from the United States (U.S.), 18% were from Europe, and 8% were from China. Of the distinct abstract authors with a known country, we were able to match 47% with age and gender. Women represented 33% of the distinct abstract authors overall and 36% of the authors from U.S., 34% from China, and 33% from the UK. Of the top 10 countries, Japan had the smallest proportion of female authors with 15%.

Not surprisingly, as the AGU Fall Meeting has always been held in the U.S., many co-author relations, 53%, are between U.S.-based authors. But still a good number of U.S. authors, including early-career researchers, have one or more international co-authors among all their connected authors.

Regardless of national affiliation countries, women had fewer international co-authors than men (Figure 4 and Table 3 in Hanson *et al.*, 2019). This is observed both across all ages and within each age cohort. The differences for each cohort indicate that having a broader overall international network is not solely an age effect (where the older average age of male authors is associated with a larger and thus more international network). About 50% of women and 60% of men have at least one international co-author (Fig 4., top). This pattern holds across major country-level data. In the U.S., for example, just over 40% of women and 50% of men have at least one international co-author. This is similar to that of China. In both U.S and China, the proportion of women in nearly all age cohorts with an international co-author is less than comparably aged men, typically by 5 to 10%, and as these countries have the largest number of authors attending the AGU Fall meeting, these dominate the overall pattern. Elsewhere, however, particularly in some countries in Europe but also Japan and Canada, a similar proportion of men and women have at least one international co-author across age groups. In the EU especially, this may reflect dynamics of science collaborations, mobility, and education across member countries.

The youngest author cohort had fewer international author groups than any other age group. The proportion of 20-29 year old men and women with international co-authors is still fairly high: 35% of U.S. women and 40% of U.S. men have at least one international collaborator; 40% and 45%, respectively, in China; and higher in Europe. Comparison with the authors where we were unable to match gender and/or age (which are generally not AGU members), but could still identify international collaborations, implies that AGU members are somewhat more connected internationally overall than non-members (46% of authors with unknown age and gender have an international connection, versus 56% of the matched authors). In data where we could match gender but not age, the same gender differences persist.

A complementary way to look at international networks is to consider the insularity of co-authors. For this analysis, we define insularity as the proportion of all co-authors in the aggregated network of authors from a country that are in the same country (Fig. 4 and Table 4 in Hanson *et al.*, 2019). For U.S. authors, about 80% of their co-authors were also in the U.S. This was the most "insular" network represented in our data and likely in part reflects the U.S. location of the meeting, AGU's membership, and the large size of the research community in the US. Japan-based authors were the second most insular with 72% of their overall network with other Japan-based authors. Authors in Switzerland (35%), the Netherlands (39%), and Spain (39%) had the least insular networks among major countries represented at the Fall Meeting. Collaborators from these countries were spread out among colleagues in the U.S., UK, and other countries in Europe. Non-U.S. authors who network with U.S. authors the most were from Canada (29% U.S.) and Africa (24%). Italy- and Japan-based authors have the smallest relative U.S. author network with 16% and 11%, respectively. China-based author networks are mostly with China (67%) and the U.S. (19%). They have relatively few connections with authors in their geographical region such as Japan (1% of their network), and other Asian countries collectively (2% of their network).

In general, younger authors' networks, especially for those in their 20s, were more insular than those in later age cohorts (Figure 4). This is not surprising because their networks are still nascent and grow out from their departmental advisor and peers. Most networks decreased in insularity as the age cohort increased.

For most countries, including the U.S. and China, women's networks were more insular than those of their male counterparts (16 of the 20 country-regions but not all are significantly different), and across most age-cohorts, consistent with the analysis above. Of these country-regions, the largest differences were in Spain (women 41% insular, men 31%), Africa (women 45% insular, men 39%), and South America (women 57% insular, men 51%). However, there are a few countries where women's networks were less or equally insular: namely Canada (women 44% insular; men 45%), Switzerland (women 31%; men 32%), and the UK (both 43%).

We separately looked for possible trends with respect to both international engagement and gender and age diversity of co-authors over the time period of the data, but many trends were not significant and none were inconsistent with these results. International attendance increased to the meeting over the years from some countries, notably China.

4. Discussion

These co-authorship patterns have several implications. First, the data are consistent with the notion that at least some of the persistent differences in invitations by men versus women for awards, reviewing, and speaker roles (among others) are related to differences in their networks. Although there could be other confounding factors, strong familiarity with colleagues via co-authorship or other close connections is a logical explanation for some of the differences. It is interesting that the observed network differences of about 5% overall and within most age cohorts is of the same magnitude as the commonly observed differences in invitations. This would also then explain why training and awareness of bias alone have not fully mitigated implicit bias: It is more difficult to overcome a real structural difference. The comparison with the overall AGU distribution suggests that this is not simply homophily (Hellmer *et al.*, 2018; Murray *et al.*, 2019)—where researchers of one gender tend to interact with or recommend each other. Women's networks are balanced with respect to the AGU gender-age distribution compared to the overall expected population whereas men's networks in the Earth and space sciences are male-dominated.

Secondly, the data also imply that women may not have the same opportunities to build international collaborations as male counterparts, or are not taking full advantage of opportunities, or are not able to take advantage of these, or all of these. Given the growing importance of international collaborations for addressing large challenges in the Earth and space sciences, and the importance of these connections for career advancement, these differences are also important to address.

Finally, differences in author networks, both with respect to gender and international diversity, are apparent among early-career scientists, and similar differences are apparent for older scientists. To the extent that early formed collaborations and networking habits persist through careers, the co-author data suggest that the "old boys club" starts as a "young boys club."

This result implies that addressing implicit bias should focus on intentionally and equitably extending collaborative opportunities to early career scientists. Programs, practices, tools and advice should be aimed at demographically extending and balancing the networks of men with respect to gender and of women with respect to international connections. Proactive efforts to promote engagement and collaborations across research groups and departments would be critical. If these network effects are socially and structurally ingrained and reinforced, as they likely are, corrective actions may take time to manifest throughout the scientific workforce.

A separate analysis comparing author groups and acceptance rates across AGU journals for the same time period (Lerback *et al.*, submitted; copy provided) shows that the acceptance rates and citations of papers with gender-diverse and internationally diverse author groups are higher than those for single-gender author teams and single-country author teams. This result in the Earth and space sciences, which supports results from other studies in other disciplines (Nielsen *et al.*, 2017; reviewed in Hall *et al.*, 2018), indicates that forming diverse teams leads to better and more impactful science. For addressing implicit bias, it implies that there is a positive feedback for creating gender diverse and internationally diverse teams—it may lead to better science. Thus, there is or should be incentives for both of these efforts. Even small benefits accrue over the course of a career and in aggregate.

Finally, this analysis emphasizes the need to account for and include age data in such analyses and the value of data within scientific societies that can be connected with their meetings and publications. We encourage other societies to work to collect this data and amplify and extend this work.

Acknowledgments: Karine Blaufuss at AGU helped greatly in accessing member data and matching records.

Data Availability: This analysis merges a public dataset on published AGU abstracts with member-provided data on year of birth, gender, and personal emails. Analysis and publication of aggregated member data are consistent with AGU's privacy policy https://www.agu.org/Privacy-Policy. Although the abstract records are public and available here: https://abstractsearch.agu.org/about/, the matched member data are covered under this policy and cannot be released publicly. Given the information contained in and used for this analysis in the abstract data and other available information, it is not possible to fully anonymize the complete merged dataset to prevent release of any member data (even if not all of it), even if much of the data set could be assembled separately from public records. Summary tables in support of the analysis are provided in this manuscript and at this Xenodo archive (DOI: 10.5281/zenodo.3445470).

References

Abramo, G., D'Angelo, C.A. & Di Costa, F. (2019), The collaboration behavior of top scientists, Scientometrics 118: 215. https://doi.org/10.1007/s11192-018-2970-9

Ford, H.L., Brick, C., Blaufuss, K. & Dekens, P.S. (2018), Gender inequity in speaking opportunities at the American Geophysical Union Fall Meeting. *Nature Communications*, 9(1), p.1358.

Fox, C.W., Paine, C.E.T. (2019) Gender differences in peer review outcomes and manuscript impact at six journals of ecology and evolution. Ecol Evol.; 9: 3599– 3619. https://doi.org/10.1002/ece3.4993

Hall, K. L., Vogel, A. L., Huang, G. C., Serrano, K. J., Rice, E. L., Tsakraklides, S. P., & Fiore, S. M. (2018), The science of team science: A review of the empirical evidence and research gaps on collaboration in science. *American Psychologist*, 73(4), 532-548. http://dx.doi.org/10.1037/amp0000319

Hanson, B., and J. Lerback (2017), Diversifying the reviewer pool, *Eos*, 98, https://doi.org/10.1029/2017EO083837.

Hanson, B., Wooden, P., Lerback, J. (2019) Datasets for Age, Gender, and International Author Networks in the Earth and Space Sciences: Implications for Addressing Implicit Bias. Xenodo http://dx.doi.org/10.5281/zenodo.3445470.

Holmes, M.A. (2015), Implicit Assumption: What It Is, How To Reduce Its Impact, in *Women in the Geosciences: Practical, Positive Practices Toward Parity*, Special Publications 70. Edited by M.A. Holmes, S. O'Connell, and K. Dutt. American Geophysical Union.

Helmer, M., Schottdorf, M., Neef, A. & Battaglia, D. (2017), Gender bias in scholarly peer review. *Elife*, 6, p.e21718.

Lerback, J. & Hanson, B. (2017). Journals invite too few women to referee. *Nature News*, 541(7638), p.455.

Lerback, J., Hanson, B., Wooden, P. (submitted to PNAS, 2019) Association between Author Diversity and Acceptance Rates and Citations in Peer-reviewed Earth Science Manuscripts (copy provided).

Murray, D., Siler, K., Lariviére, V., Chan, W.M., Collings, A.M., Raymond, J. & Sugimoto, C.R., (2019). Author-Reviewer Homophily in Peer Review, *bioRxiv* 400515, https://doi.org/10.1101/400515

Nielsen, M. W., Alegria, S., Börjeson, L., Etzkowitz, H., Falk-Krzesinski, H. J., Joshi, A., Leahey, E., Smith-Doerr, L., Williams Woolley, A., & Schiebinger, L., (2017), Opinion: Gender diversity leads to better science, *Proc. Nat. Acad. Sci.*, 114 (8) 1740-1742; https://doi.org/10.1073/pnas.1700616114

Roper, R.L. (2019) Does Gender Bias Still Affect Women in Science? *Microbiology and Molecular Biology Reviews* 83 (3) e00018-19; DOI: 10.1128/MMBR.00018-19

Teplitskiy, M., Acuna, D., Elamrani-Raoult, A., Körding, K. & Evans, J. (2018), The sociology of scientific validity: How professional networks shape judgement in peer review. *Research Policy*, 47(9), pp.1825-1841.

US Bureau of Labor Statistics. *Women in the Laborforce: A Databook Report 1059* (2015).

West, J.D., Jacquet, J., King, M.M., Correll, S. J., & Bergstrom, C. T., (2013), The Role of Gender in Scholarly Authorship, *PLOSOne* https://doi.org/10.1371/journal.pone.0066212

Wu, L., Wang, D., & Evans, J. A. (2019), Large teams develop and small teams disrupt science and technology, *Nature* 566, pages 378–382

Figure Captions:

Figure 1. cumulative author networks (in percent of total population) by age and gender for each age-gender cohort. Horizontal hashes indicate the age-gender distribution of AGU members over this time period..

Figure 2. Percentage of women within each age cohort in the overall co-author network of men and women of each age cohort. Horizontal hashes indicate the age-gender distribution of AGU members over this time period. Error bars are +/- 95% standard confidence intervals.

Figure 3. (top) Number of abstracts per per person, acting as any co-author, by gender and age cohort. (bottom) Average number of co-authors by age cohort and gender.

Figure 4. International author networks. (top) Percentage of authors by age and gender with at least one international co-author. (bottom) Insularity, or the percent of co-authors of the aggregated author network that are all within the same country.

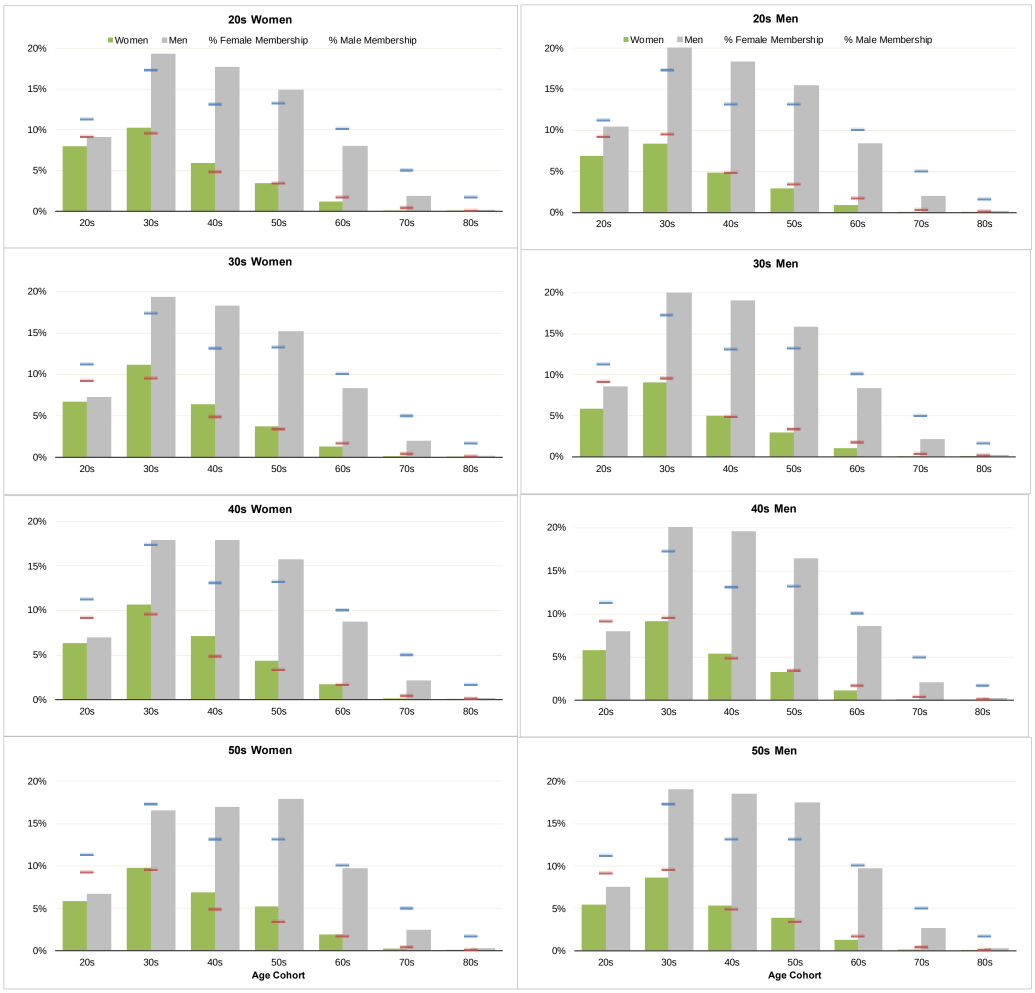

20s Women
Women
Men
% Female Membership
% Male Membership
20s Men
30s Women
30s Men
40s Women
40s Men
50s Women
50s Men
0%
5%
10%
15%
20%
20s
30s
40s
50s
60s
70s
80s
Age Cohort

Figure 1.

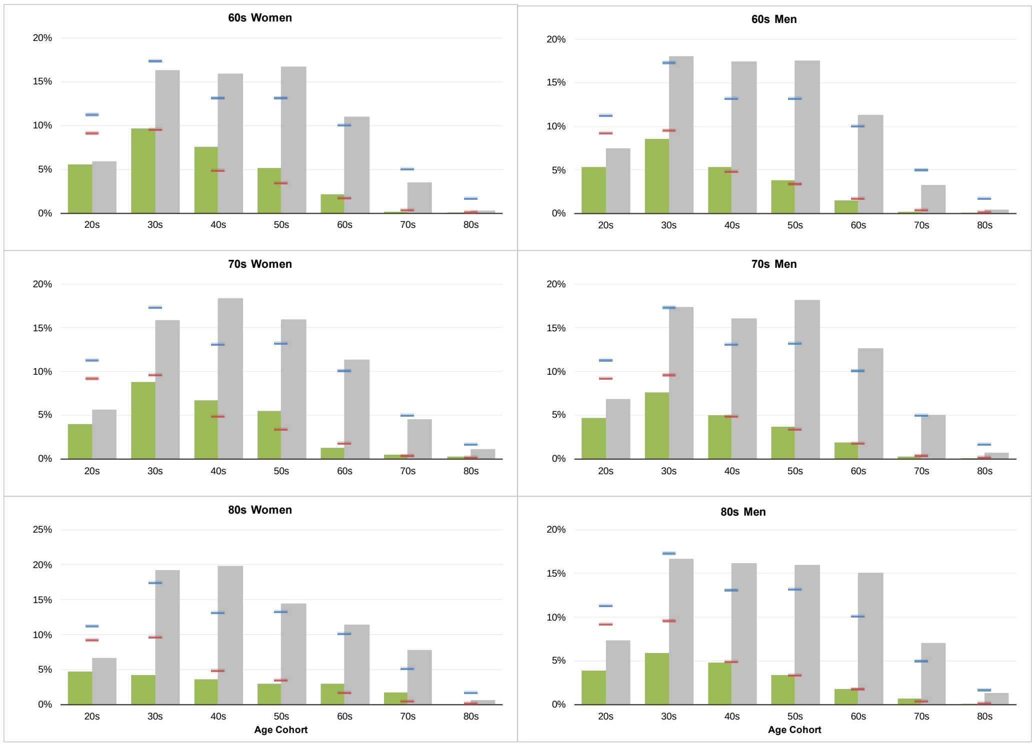


Figure 1 Continued

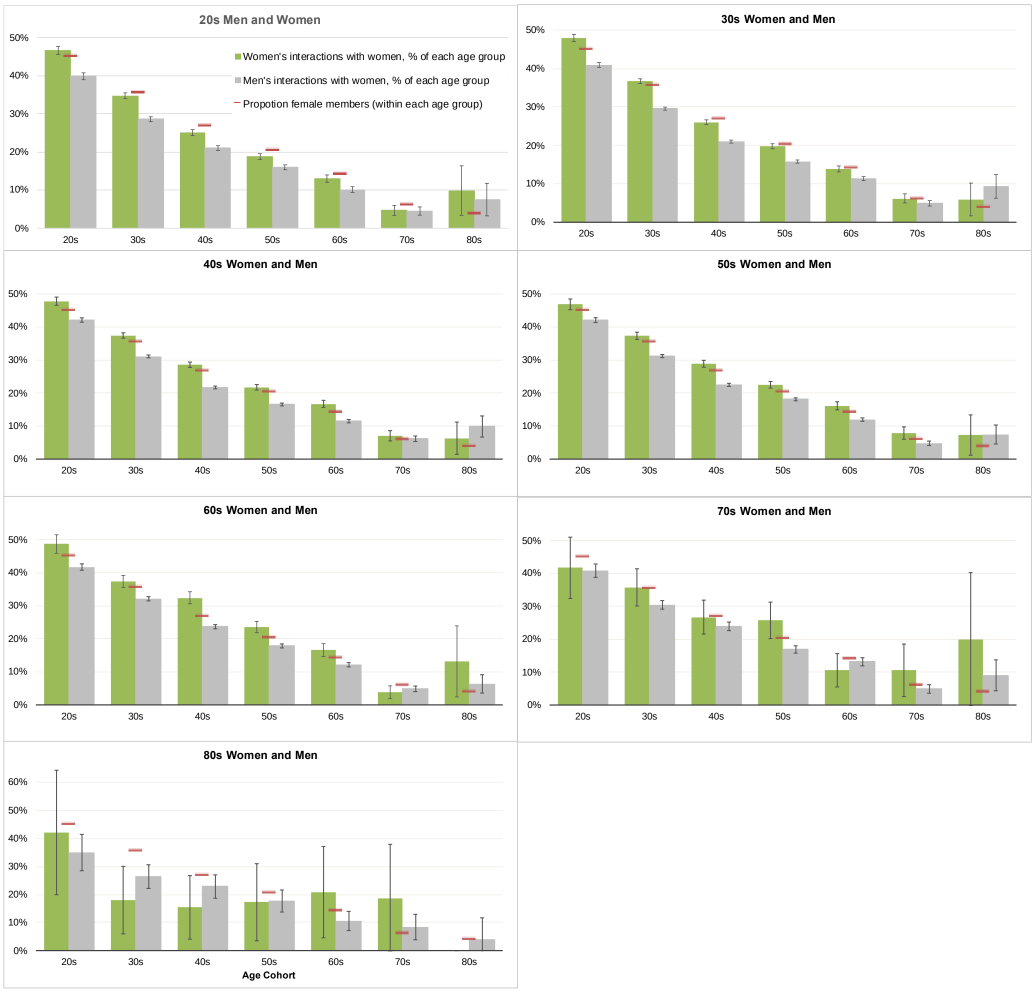

20s Men and Women
Women's interactions with women, % of each age group
Men's interactions with women, % of each age group
Propotion female members (within each age group)
30s Women and Men
40s Women and Men
50s Women and Men
60s Women and Men
70s Women and Men
80s Women and Men
0%
10%
20%
30%
40%
50%
60%
20s
30s
40s
50s
60s
70s
80s
Age Cohort

Figure 2

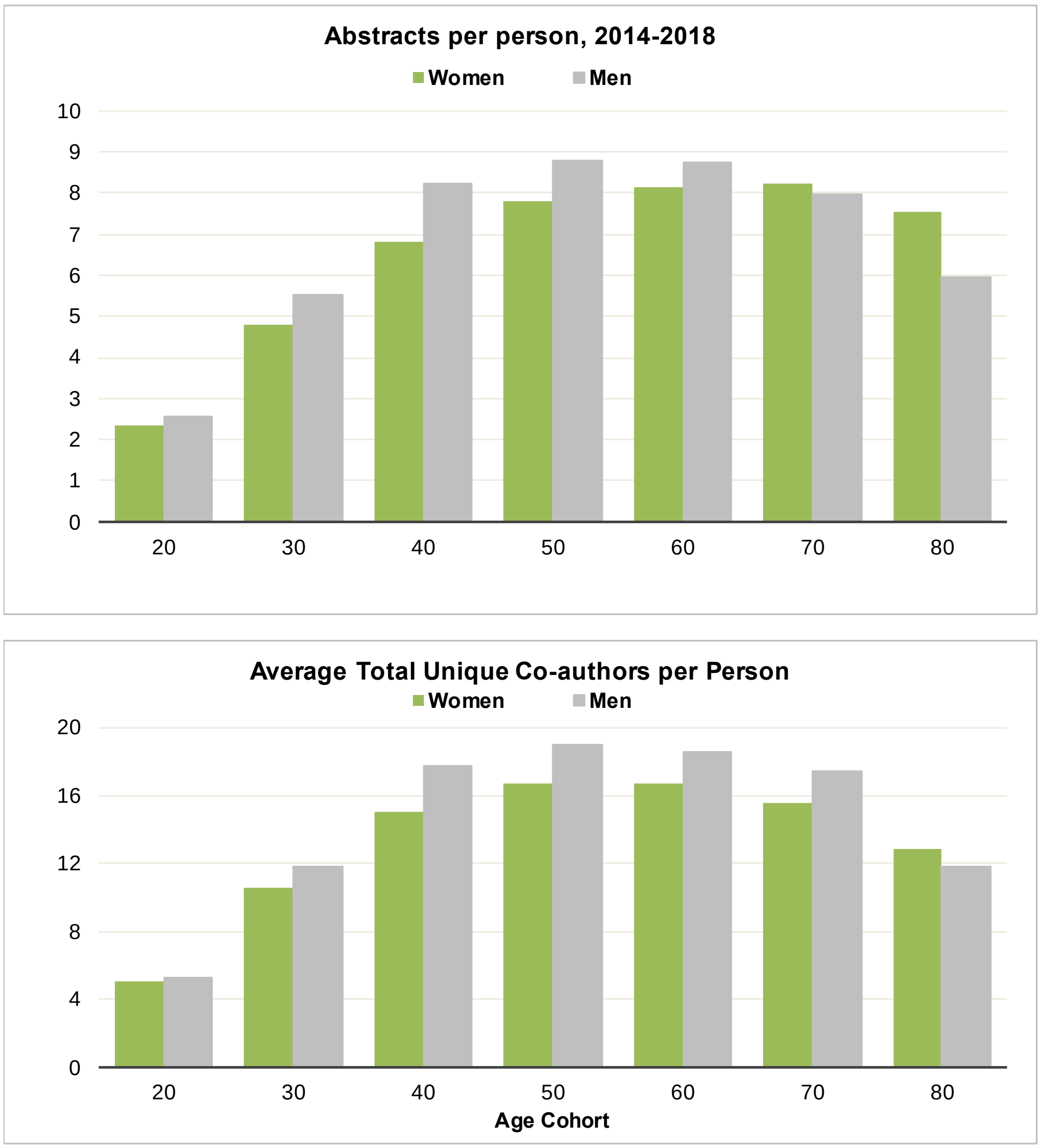
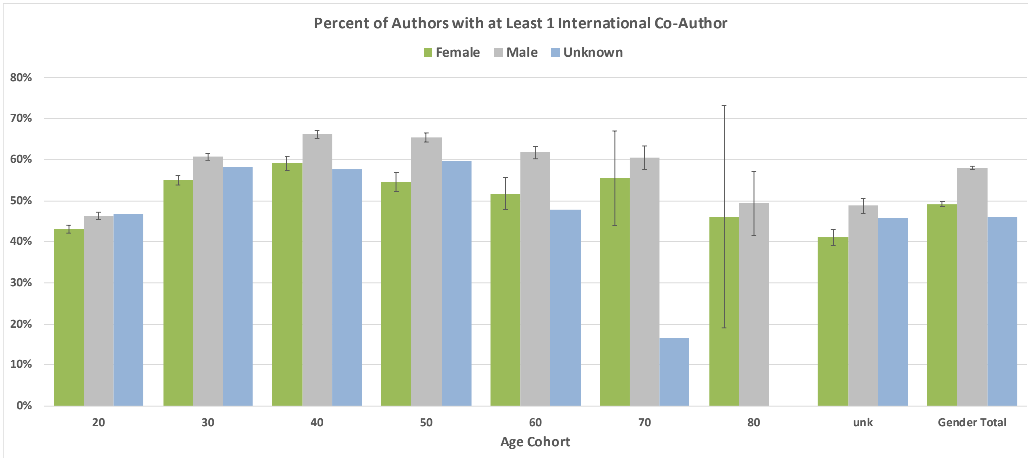

Abstracts per person, 2014-2018
Women
Men
10
9
8
7
6
5
4
3
2
1
0
20
30
40
50
60
70
80
Average Total Unique Co-authors per Person
Women
Men
20
16
12
8
4
0
20
30
40
50
60
70
80
Age Cohort

Figure 3.


Percent of Authors with at Least 1 International Co-Author
Female
Male
Unknown
80%
70%
60%
50%
40%
30%
20%
10%
0%
20
30
40
50
60
70
80
unk
Gender Total
Age Cohort

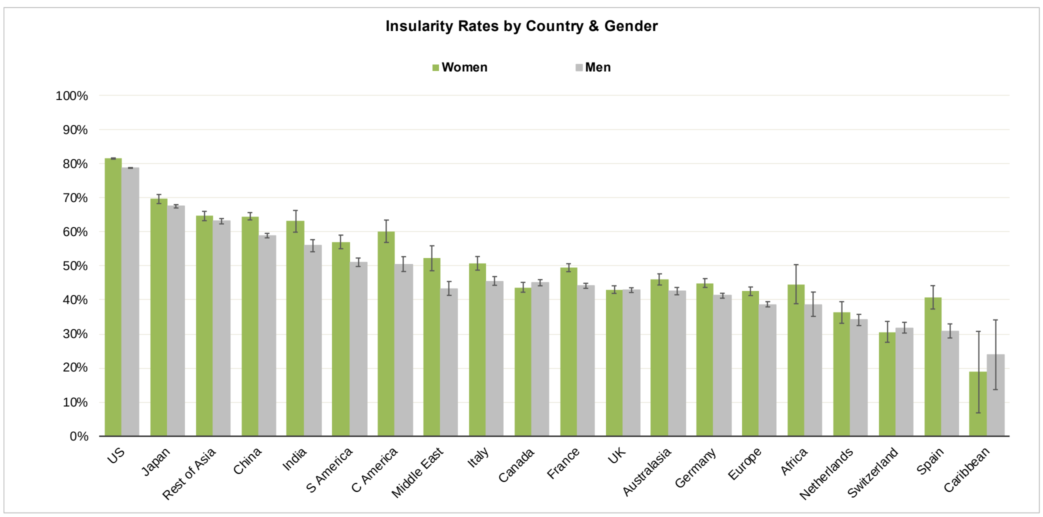

Insularity Rates by Country & Gender
Women
Men
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
US
Japan
Rest of Asia
China
India
S America
C America
Middle East
Italy
Canada
France
UK
Australasia
Germany
Europe
Africa
Netherlands
Switzerland
Spain
Caribbean

Figure 4.