Open science and communal culture promote women's participation, diversity, and discovery

Author block 1:
Mary C. Murphy*, Amanda Mejia*, Jorge Mejia*, Xiaoran Yan*,

Author block 2:
Sapna Cheryan, Nilanjana Dasgupta, Mesmin Destin, Stephanie A Fryberg, Julie A. Garcia, Elizabeth L. Haines, Judith M. Harackiewicz, Alison Ledgerwood, Corinne A. Moss-Racusin, Lora E Park, Sylvia, P., Perry, Kate A. Ratliff, Aneeta Rattan, Diana T. Sanchez, Krishna Savani, Denise Sekaquaptewa, Jessi L. Smith, Valerie Jones Taylor, Dustin B. Thoman, Daryl A. Wout,

Author block 3:
Patricia L. Mabry**, Susanne Ressl**, Amanda Diekman**, Franco Pestilli**

* shared first author contribution
** shared senior author contribution

ABSTRACT

Science is undergoing rapid change, with the movement to improve science focused largely on reproducibility and open science practices. This movement offers the opportunity to recreate a scientific culture in a more inclusive model. The culture of science has long valued and rewarded individual brilliance and competitive winner-take-all models, though much of scientific work is actually interdependent and collaborative. Collaborative, interdependent cultures can attract women and other underrepresented groups and catalyze discovery. At the same time, more scientists and funding agencies are recognizing the value of multidisciplinary "team science." Yet the movement to improve science does not necessarily guarantee greater diversity: a communal culture may be essential for reaping the potential benefits of social diversity for scientific progress. We examined the open science and reproducibility literature across all scientific disciplines to investigate the emerging cultures in the movement to improve science. We collected a dataset of approximately 3,000 journal articles and conference proceedings published between 2000 and 2017 in these literatures. We found that these approaches have few common papers or authors, suggesting that open science and reproducibility literatures are emerging relatively independently. Network analyses revealed that open science literature has a more collaborative structure, where a greater proportion of papers share authors. Semantic analyses of paper abstracts revealed that open science and reproducibility literatures appear to be adopting different cultural practices and frames, namely that the open science literature includes more language reflecting communality and prosociality. Finally, consistent with the literature suggesting the diversity benefits of open, communal and prosocial purposes, we found that women publish more frequently in high-status author positions (first or last) within open science compared to reproducibility. We conclude with a list of actionable suggestions for cultivating a collaborative and diverse culture of science.

Author Contributions. Literature search: MCM, ABD, PLM, FP, CRS, EH, AL, KR, ND, DT, LP, JS, SP, JH, KS, VJT, AR. SC. SF, CAM-R, DW. AL, DS, JAG, DTS . Figures: XY, AM, PLM, SR Study design: MCM, PLM, ABD, SR, FP, AM, JMM. Data Collection: XY, ELH, AL, KAR, JMM. Data analyses: AM, XY, JMM, PLM. Data interpretation: AM, XY, JMM, PLM, MCM, FP. Writing: MCM, FP, SR, PLM, ABD, AM, CRS, ELH, AL, KAR, ND, DT, LP, JLS, SPP, JMH, KS, VJ-T, AR, SF, MD, AR, SAF. Effectively three groups of authors contributed to this work with different degrees of contribution. Group 1 first authors. MCM, AM, JM, XY. Group 3 senior contributors. PLM, SR, AD, and FP. Group 2 study supporters.

Acknowledgements. M.C.M. was supported by NSF CAREER DRL-1450755 and NSF HRD-1661004. A.B.D. was supported by NSF GSE-1232364. F.P. was supported NSF IIS-1636893, NSF BCS-1734853, NSF AOC-1916518, NIH NCATS UL1TR002529, a Microsoft Research Award, a Google Cloud Award, the Indiana University Emergent Area of Research Initiative “Learning: Brains, Machines, Children” and Pervasive Technology Institute. The authors would like to acknowledge Angela Sharpe, Cassidy Sugimoto, for her thoughtful discussion about the topics in the paper which were reflected in the manuscript, and Michael Jackson for help with the design of the figures.

INTRODUCTION

At the current moment, science is undergoing a "revolution" to better itself (1). The aim of this revolution is bold. At its core, the movement to improve science encompasses two goals: (a) understanding the flaws, weaknesses, and reproducibility of past scientific processes and findings, and (b) improving research practices through greater rigor and transparency (e.g., open sharing of data, code, resources; standardized statistical procedures; pre-registration). As with any revolution, a time of unrest can also be a time of opportunity. In this article, we examine the emerging cultures in the movement to improve science, and investigate the representation of women in these emerging subcultures.

In cultural analyses, the actions and cognitions of individuals stem from and produce the norms and practices of groups and institutions (2). The emerging subcultures of the current scientific reform movement have their roots in the broader culture of science, technology, engineering, and math (STEM) that can serve as a barrier to the inclusion and advancement of women (3–5). The culture of science has long valued individual brilliance, competition, and a winner-take-all model of success (6). In particular, people inside and outside of STEM perceive STEM fields as affording more opportunities for individual success and achievement than for prosociality and collaboration (7). The scientific practice of rewarding individual achievement has perhaps unwittingly fostered a more independent, competitive culture that ignores and possibly even disincentivizes cooperation (8, 9). These cultural practices have implications for who joins and advances within scientific fields: the perceived lack of prosocial and collaborative culture in STEM has been shown to deter women especially (7, 10). Indeed, the presence of collaborative practices and prosocial purposes may be particularly important in fields focused on scientific reform: Critiquing established authors or practices, no matter how well intended or delicately stated is often interpreted as criticism and puts the critic in a defensive position. The role of critic may be particularly risky and unappealing to female scientists. First, women may feel less able to voice dissent (particularly when solo) against established figures, because this conflict-prone stance both violates gender role expectations (11) and their minority status in most scientific fields ((12) and NSF, 2017). Second, these challenges to the establishment may be construed as for the benefit of the individual (i.e., gaining recognition) rather than for the collective good (i.e., to improve and advance science).

The movement to improve science, to date, can be characterized by two contrasting motifs. One focus has been to assess the reproducibility of previously published scientific results (13). A second focus is to develop open science practices that facilitate the sharing and reuse of research assets (e.g., data, code) in order to improve rigor and accelerate the rate of scientific discovery (14–17). For shorthand we refer to these two literatures as "reproducibility" in the first case, and "open science" in the second. In this article, we examine whether these two approaches to the improvement of science have embodied different cultures in their (a) interconnectedness and (b) prosocial focus, and whether women's participation differs across these approaches. We also discuss the implications of our findings with respect to the impact of these cultures on women in science and vice versa.

Researchers involved in the efforts to improve science have acknowledged a gender diversity problem (18, 19), and this time of reform offers the opportunity to reinvent scientific culture in a more inclusive mode, which, as we will discuss, may in fact be better for science. If the movement to improve science perpetuates a scientific culture that prioritizes independent, dominant, or adversarial values, we risk leaving many talented women at the margins, feeling unwelcome and excluded (20).

Here we explored whether the reproducibility and open science literatures–both of which share the goal of improving science–exhibit different (a) collaborative structures, (b) explicitly prosocial foci, and (c) engagement of female scientists. Our focus on author gender was motivated, in part, by the ability to leverage existing archives of scientific publications, but we would predict similar findings for authors from other underrepresented groups. We anticipated that the reproducibility vs. open science literatures would have different levels and patterns of participation by women associated with the different collaborative structures and cultural frames adopted within each literature. When fields are more adversarial, individuals from underrepresented groups (such as women) may be less motivated to engage (21). In contrast, fields that emphasize collaborative or prosocial norms may inspire greater participation among underrepresented groups (22).

Our team conducted network analyses of the open science and reproducibility literatures and found that these literatures have few common authors–suggesting these approaches have developed relatively independently from each other. Moreover, text analyses of article abstracts revealed that the open science and reproducibility literatures appear to be adopting different explicit cultural frames. Open science includes significantly more language that reflects the cultural values of communality and prosociality. We also find that the contributions of women in these literatures differ. Women, for example, are more likely to occupy high-status author positions (taking the first or last position in published papers) within the open science literature than in the reproducibility literature (see Figure 3).

Taken together, we find that despite its current biases (18), the open science focus of the movement to improve science has the seed of an interconnected and prosocial culture that, if cultivated and embraced, may attract more women and underrepresented people in general. We believe that the collaborative, forward-looking aspect of open science has the potential to facilitate greater diversity and inclusiveness. Like all of science, investigators interested in the improvement of science should nurture a culture that attracts and retains a diversity of people (23–25).

RESULTS

We performed both network science and semantic text analyses to establish the structural landscape and cultural foci of the open science and reproducibility literatures and women's participation in them. To do so, our team analyzed data from Microsoft Academic Graph (26), consisting of 2,926 scientific articles and conference proceedings (hereafter referred to as "papers") published between 2010 and 2017 that included "Open Science" or "Reproducibility" as a field of study (see **Materials and Methods** and **Supplementary**

Information). This sample consisted of 879 Open Science papers and 2,047 Reproducibility papers.

Differences in authors' community structure in the open science and reproducibility. We analyzed a total of 3,157 unique article author IDs in the Open Science literature and 8,766 in the Reproducibility literature. We built two collaboration networks using these author names (**Figure 1**). Nodes in these networks represent scientific articles, two nodes share an edge if at least one author appears in both papers (see **Methods** for details). Results revealed that the Open Science network contained 879 nodes and 389 edges, while the Reproducibility network contained 2,047 nodes and 856 edges. Importantly, the Open Science network is more edge dense (0.101%) than the Reproducibility network (0.041%), demonstrating a higher degree of collaboration in the Open Science literature (one-sided Fisher's exact test $p < 0.001$). We also performed a "connected components analysis" of the network (27, 28) to measure the degree of isolation of individual sub-networks of authors within each literature (see **Methods**). Results show that the Reproducibility network contains more isolated author networks than the Open Science network with 1,641 versus 661 components, respectively (we note that a higher number of components indicates that the network is more disconnected). Accounting the size differences of the two networks, we show that the average component size (ACS) is also higher for the Open Science network (ACS: 1.33 vs. 1.25). **Figure 1** visualizes the two networks to facilitate interpretation of the observed differences between the two kinds of literature in terms of network connectedness and fragmentation. In sum, the Open Science literature was found to have a greater number of connections (shared authors) between papers and the Reproducibility literature contains more isolated author networks – and these differences between the two kinds of literature are statistically significant ($p < 0.01$ as reported above).

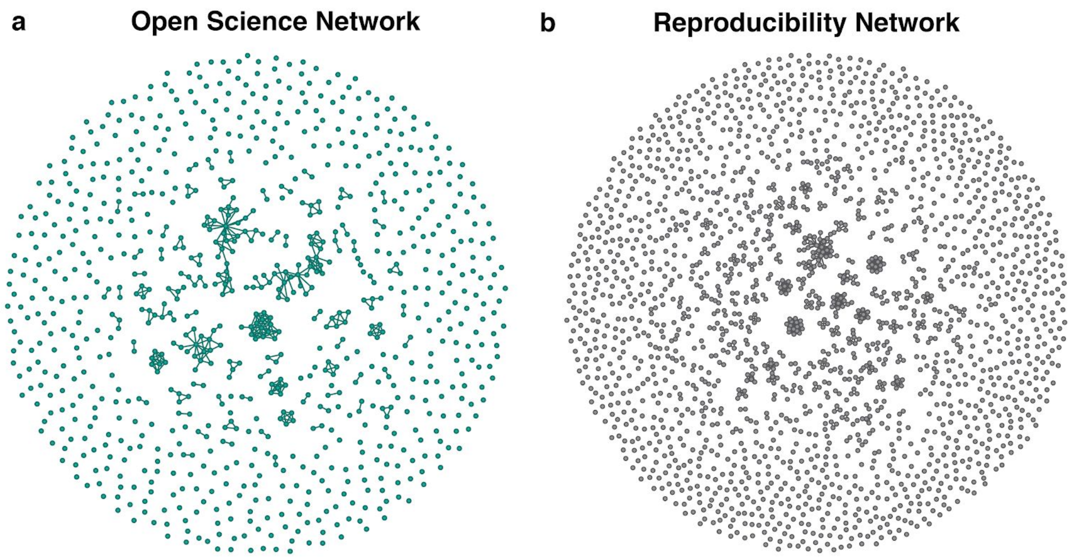

Figure 1. Differences in author community structure: Open Science vs. Reproducibility. Each circle, or node, represents a scientific article. Articles share an edge (line connecting two nodes) if at least one author appears in both papers. While networks in both literatures are relatively sparse, the open science literature has formed a larger collaboration network (i.e., the community structure or a group of highly connected nodes in the center of the visualization), when compared to the reproducibility network. Data visualized using Gephi (29).

Women are more likely to be represented in high-status author positions in the open science. Women scholars are significantly more likely to be represented in high-status author positions (i.e., first or last author position) in the open science literature than in the reproducibility literature. **Figure 2** displays gender representation for the open science and reproducibility literatures for single- and multi-authored papers. The single-authored subset includes 255 Open Science papers and 342 Reproducibility papers, while the multi-authored subset includes 624 Open Science papers and 1,705 Reproducibility papers. Due to different field conventions, we consider a scholar to hold a high-status authorship position if they occupy *either* the first or last author position within a multi-authored paper. We find that, overall, women are significantly less likely than men to publish single-author papers in both literatures, based on papers with identifiable author gender. An exact one-sided Binomial test indicated that the percentage of female single authors (among single-author papers with identifiable author gender) of 33.0% in the Open Science literature and 28.1% in the Reproducibility literature was lower than the expected proportion of 50% ($p < 0.001$ for both tests). For the remaining analyses we focus on multi-author papers.

Women hold high-status authorship positions in 60.6% of the multiple-author papers in the Open Science literature, compared to 57.9% in the Reproducibility literature (among multi-author papers where female lead authorship could be determined). Note that if women and men were equally represented in first and last author positions, the expected percentage of multiple-author papers with a female in a high-status (first or last) author position would be 75% (comprised of a 25% chance of female first and last, a 25% chance of female-first and male-last, and a 25% chance of male-first and female-last).

We performed a regression analysis to better understand gender differences in high-status authorship positions across the two literatures. Specifically, we fit a logistic spline regression model controlling for time trends, team size, and manuscript type (i.e., journal article, conference proceedings). For this analysis, we used the subset of multi-authored papers for which we were able to conclude whether or not a woman holds a high-status position (i.e., where with some degree of confidence, the gender of the first and last author could be determined, or the gender of the first or last author could be identified as female even if the other could not be identified). We also excluded 28 Open Science papers and 40 Reproducibility papers with more than 12 authors to avoid giving these papers disproportionate influence on the regression fit. The resulting dataset consisted of 454 Open Science papers and 955 Reproducibility papers. After controlling for team size, year of publication and manuscript type, we found that multi-author papers in the Reproducibility literature have 61% lower odds of having a woman in a high-status authorship position compared with the Open Science literature ($p<0.001$; see **Supplementary Information Table S1**). Thus, whereas women are underrepresented in high-status author positions in both literatures, there is significantly greater representation of female authors in high-status author positions in the Open Science literature than in the Reproducibility literature. Additionally, we fit the model controlling for the academic field of study (i.e., mathematics, biology, medicine) and found similar effects.

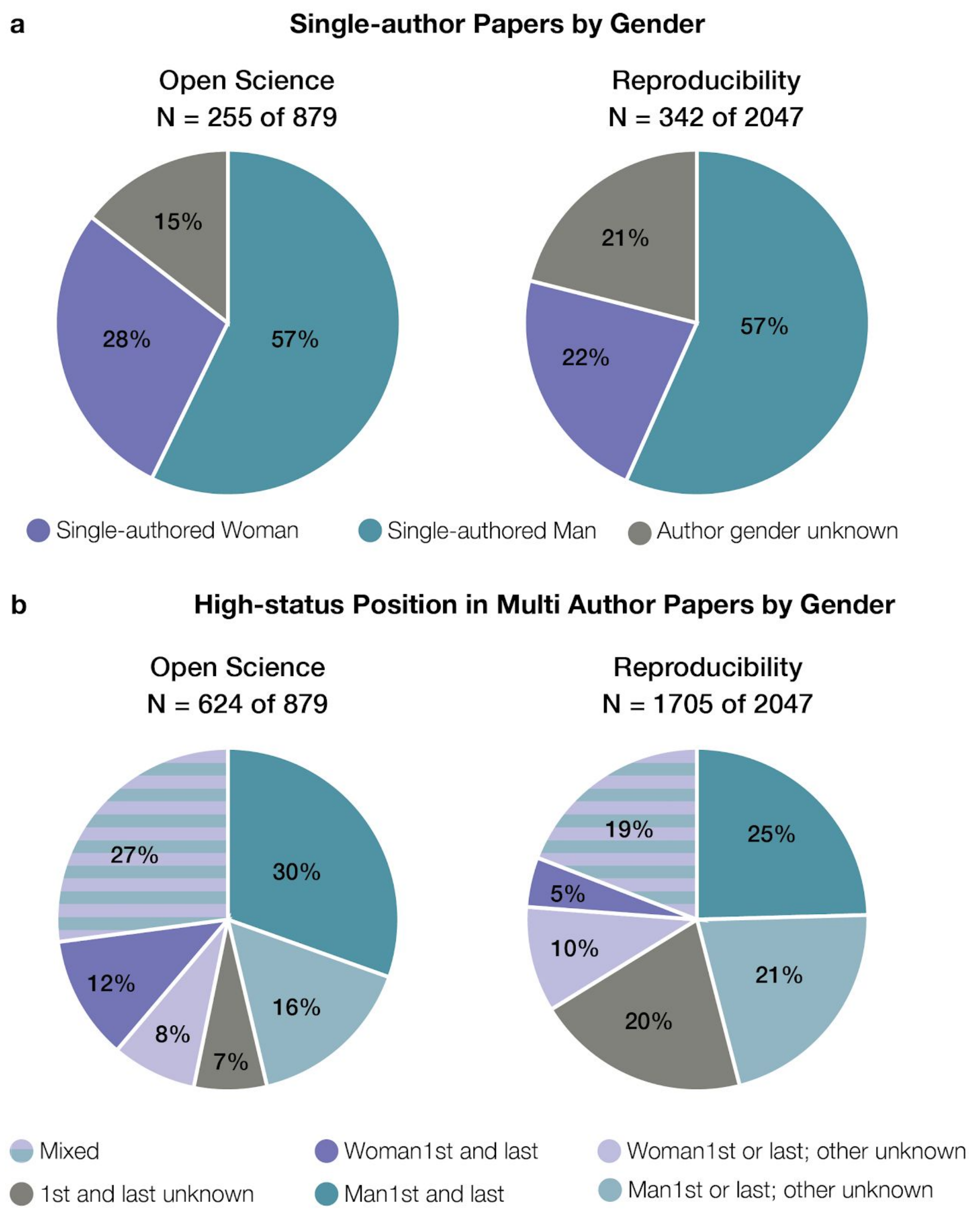

Figure 2. Gender representation in high-status author positions (first or last) in open science and reproducibility. a. Single author paper by gender. Women are underrepresented in single-authored papers in both the open science and reproducibility literatures. **b. High-status position in multi author papers by gender.** Women are underrepresented in high-status author positions in both literatures, but have greater representation in Open Science (with 47% with known female first *or* last author and 12% with known female first *and* last author) than in the Reproducibility literature (with only 34% with known female first *or* last author and only 5% with known female first *and* last author).

Team size in multi-author papers is associated with women's leadership in open science and reproducibility. The distribution of team size within each literature illustrates that women's likelihood of authoring in high-status positions in multi-author papers (i.e., first or last author) within the Open Science literature is greatest in smaller teams (2-3 author papers) and remains consistent as teams become larger (see **Figure 3**). However, within the Reproducibility literature women are relatively less likely to author in high-status positions in small teams (2-3 author papers). The results of regression analysis confirm this difference after controlling for other important variables, including publication year and manuscript type (see **Figure 4**, left panel). One interpretation of this finding is that it may be more threatening for women to lead smaller-team articles in the Reproducibility area.

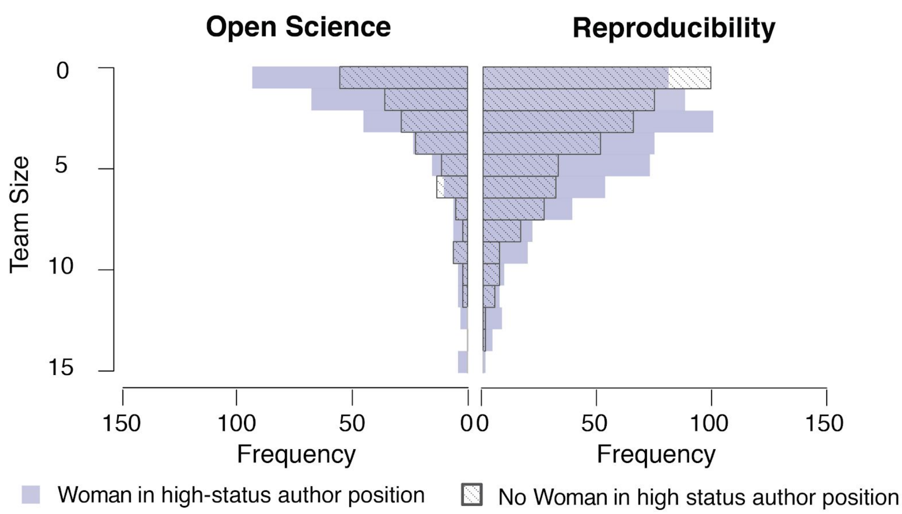

Figure 3. Team size and women's representation in high-status positions in multi-author papers. Women's representation in high-status authorship positions (first and last authorship) is affected differently by team size in the open science and reproducibility literatures. Women more frequently assume high-status positions in both smaller and larger teams in Open Science, while they do so in only in larger teams in the Reproducibility literature.

Women's participation over time is increasing in open science and decreasing in reproducibility. Our regression analysis reveals that in the Open Science literature, the representation of women in high-status authorship positions (first or last author) has grown over time, while it has declined or failed to increase in the Reproducibility literature. We find that the odds of a woman holding a high-status position in the Open Science literature has grown at a rate of approximately 15.6% ($p<0.01$) year-over-year from 2010 to 2017 (see **Supplementary Information Table S1**), controlling for team size and manuscript type. In the Reproducibility literature, over the same time period the representation of women in high-status positions has *declined* at an estimated rate of approximately 3.6%, though this is not statistically significant ($p=0.20$). Examining the difference between these slopes reveals a statistically significant difference between women's representation over time between these literatures ($p<0.01$). **Figure 4** (right panel) illustrates the difference in trends over time between the two literatures on the probability scale.

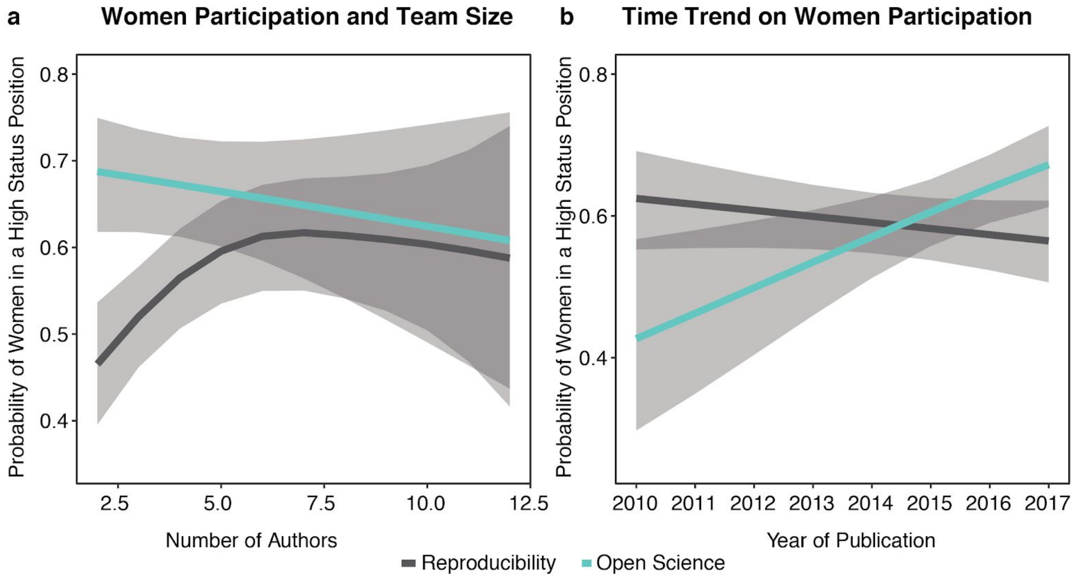

Figure 4. Estimated regression effects of team size and year of publication on women's representation in high-status positions in multi-author papers. a. Women participation and team size. Women have higher rates of high-status authorship in larger teams within Reproducibility, while rates are comparatively and consistently high in Open Science across team sizes. **b. Women participation over time.** In Open Science, the representation of women in high-status positions has grown over time, while in Reproducibility it has declined. Values are logistic regression estimates shown on the probability scale, with 95% confidence intervals indicated in grey. To produce the estimates, the x-axis variable and literature category are varied, while the remaining model variables are fixed (publication year = 2017; team size = 4 (near the mean value); and manuscript type = journal.

Text analyses suggest that the explicit culture of the open science and reproducibility literatures are different. Using a validated text-mining dictionary (30), we measure the presence of communal and prosocial constructs (e.g., contribute, encourage, help, nurture; see **Supplementary Table S3** for the list of constructs used) in the abstracts of the papers from both literatures. We excluded papers with no available abstract and those with non-English titles. The resulting dataset included 595 Open Science papers and 1,169 Reproducibility papers. In the Open Science dataset, 76% of the articles used words associated with communal and prosocial constructs, whereas in the Reproducibility dataset only 44% of the articles did (2-sided test for equality of binomial proportions, $p < 0.001$). We computed the *prosocial word density* within each dataset as the percentage of words in each abstract that reflect communal and prosocial constructs (**Figure 5** see **Materials and Methods**). The Open Science abstracts included more communal and prosocial words than the Reproducibility abstracts (Open Science mean density of 2.4%, median density 1.8%; Reproducibility mean density 0.9%, median density 0.0%). A two-sided permutation test for differences in the mean and median scores in each dataset shows that the Open Science literature includes significantly more frequent use of communal and prosocial words than does the Reproducibility literature (p<0.001 for mean and median).

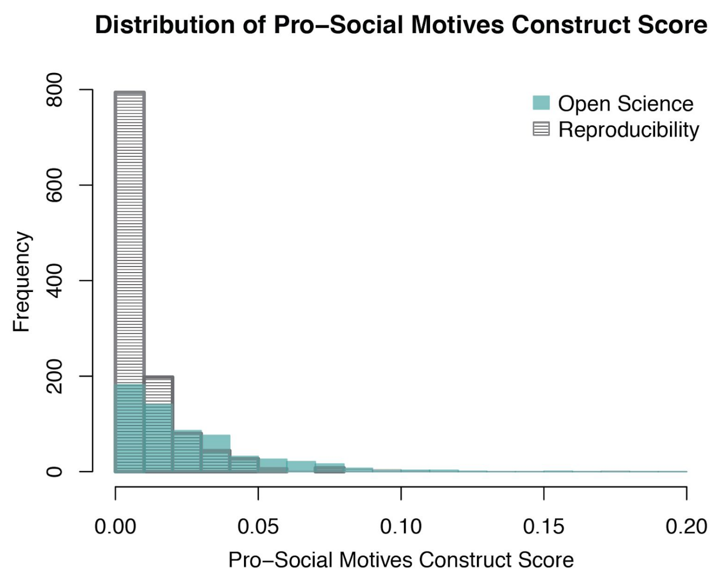

Figure 5. Distribution of communal and pro-social word density of abstracts in the Open Science and Reproducibility literatures*.* Abstracts in the Open Science literature include significantly more words associated with communality and prosociality than those in the Reproducibility literature.

DISCUSSION

Our results reveal that the movement to improve science consists of two relatively independent groups with differing approaches: (a) Open Science and (b) Reproducibility. These literatures have relatively few common authors and show significantly different community structures of how authors contribute to individual papers. Whereas the Open Science literature is significantly more interconnected with respect to co-authorship, the Reproducibility literature contains more distinct groupings of coauthors. We find that women are more likely to be represented in high-status author positions in the Open Science literature than in the Reproducibility literature. Interestingly, team size seems to be associated with women's leadership positions in both Open Science and Reproducibility, suggesting that a more collaborative scientific model may promote women's participation. Our analysis of women's participation over time reveals that it is increasing in Open Science and decreasing in Reproducibility. Lastly, our semantic text analyses suggest that the nature of explicitly prosocial cultures in the Open Science and Reproducibility literatures differ. Open Science abstracts include more explicitly communal and prosocial terms than do Reproducibility abstracts.

Of course, one problem with an adversarial, non-communal scientific culture is that it does not reflect how scientific work actually unfolds–particularly with today's emphasis on grand challenges, transdisciplinary investigations, and network science. Indeed the (false) prototype of a scientist is one in which an individual scientist (usually a white male) toils away alone in his laboratory until a flash of insight occurs in a eureka! moment (31–33). This culture is epitomized by some of our most prestigious awards that celebrate individual efforts and contributions over that of teams (e.g., Nobel prize, MacArthur Genius Award, NIH Director's Pioneer Award, NSF Career Award; NIH "independent investigator" categorization). Moreover, faculty evaluations for tenure and promotion continue to prize individual performance almost exclusively--in some cases requiring scientists to show their independent contribution to collaborative projects and/or calculating the number of first- or last-authored (vs. co-authored) publications (34, 35). Today's science relies on teams coordinating their efforts to share insights and methods, build on past work, and develop new questions and approaches (36). These collaborative and complementary processes occur locally (e.g., direct work with other laboratories) as well as globally (e.g., broadening the scientific community, sharing equipment, data, and access (37)). Science today is more likely to be a collaborative than individual endeavor, where team size can matter. Indeed, larger and more diverse teams may be necessary to realize higher impact (38). A second problem suggested by these findings is that non-communal practices and values may deter people who value communal, interdependent, and prosocial goals, including women (22), underrepresented minorities (23), first-generation college students (24), and communally-oriented men (22). If the movement to improve science is to harness this diversity, the Open Science focus currently appears to be more welcoming and inclusive than Reproducibility.

Relatively recently (in the long history of science), a confluence of factors–including the development of the Internet and associated collaborative technologies along with evidence for the value of team-based science–has enabled teams of scientists to collaborate in ways that were not possible before. A problem, however, is that while science today is increasingly team-based, homophily processes mean that many of these teams are likely to be relatively homogeneous with regard to sociodemographic, behavioral, and intrapersonal characteristics (39). During this time of transition away from solo work or small, siloed teams, we are witnessing a culture clash across the disciplines between the traditional culture of science–focused on competition and individualism–and the new culture of collaboration, cooperation, and sharing across the disciplines and across the globe (40).

Indeed, there is an increasing appreciation among scientists and funding agencies that multidisciplinary "team science" is required to tackle the most pressing scientific, social, and health problems of our times. Over the last decade, organizations including NIH, NSF and others have dedicated resources to facilitating team science. This work is evidenced by interdisciplinary and multidisciplinary team requirements in federal funding announcements and programs (e.g., NIGMS Collaborative Program Grant for Multidisciplinary Teams, NSF Office of Multidisciplinary Activities NIH Interdisciplinary Program in the Common Fund and its predecessors in the NIH Roadmap, NCI's Science of Team Science Toolkit, NSF BIGDATA HUB Program, NSF Collaborative Computational Neuroscience Program, NSF Office of Multidisciplinary Activities) and many other programs under the NSF and NIH roadmaps and priorities). Moreover, funders are actively attempting to address the underrepresentation of women and minorities (e.g., NSF Broadening Participation), though there are still inequities in these processes (41).

The complexity of the problems we are now facing in science demand the expertise of multiple disciplines working in coordinated fashion (42, 43). For example, addressing the problem of opioid addiction requires the integrated knowledge of researchers who specialize in pain, addiction, neuroscience, economics, computer science, psychology, sociology, biochemistry, demography, medicine, and public health, just to name a few. Intellectually diverse, multidisciplinary teams create new insights by combining existing knowledge in innovative ways (44, 45). In fact, data from the U.S. Patent and Trademarks Office show that patents generated by teams represented more breakthroughs, landing among the top 95% of all cited patents, than those from lone inventors, suggesting their generative contributions (46). Similarly, multi-authored articles are more often cited than single-authored articles (38, 47, 48) and while some have argued that this could be due to self-citation, others have suggested that it is more likely that highly collaborative projects include more diverse data and higher quality ideas, which result in greater impact (49). Importantly, it has also been suggested that whereas large teams advance science and technology, small teams can disrupt the established scientific understanding. Both types of contributions seem to be of fundamental importance (50, 51).

Establishing a communal culture can increase diverse participation: an example from structural biology. Structural biology has yielded to several female Nobel Prize awardees in STEM: Marie Curie (1911), Irène Joliot-Curie (1935), Dorothy Hodgkin (1964), Ada E. Yonath (2009). Female structure biologists were and are crucial contributors to the rise of protein structure determination. A few notables include Dame Louise Johnson, Eleanor Dodson, Elspeth Garman, Bridget Carragher, Elena Orlova, Eva Nogales, and reflect the egalitarian and collaborative culture of the structure biology community (52). Structure biology thrives from collaborative research across scientific disciplines and within the community. The Collaborative Computational Project No. 4 Software for Macromolecular X-Ray Crystallography (CCP4, ccp4.ac.uk) and the Collaborative Computational Project for Electron cryo-Microscopy (CCP4em, ccpem.ac.uk) are two examples that forged the path for Open Science. This software developer community realized early on that structural biology would benefit from open and shared science. To facilitate these aspirations, open software platforms, code sharing, standardization of file formats, and creation of data banks were needed. Recourse platforms such as the structure data bank (rcsb.org) and "Unified Data Resource for 3-Dimensional Electron Microscopy (EMDatabank, emdatabank.org) are specific examples. Reproducibility became a focus point of the structural biology research community following the retraction of 5 articles in the journal *Nature* (53). In the aftermath, the community pushed for raw data submission, evaluation tools and standards to increase reproducibility and check quality of protein structures (54). While not perfect (55, 56), the structural biology community -- with its established broader participation of women and open culture -- is a success story, demonstrating that diversity, openness and team science can promote scientific progress, as evidenced by the changes designed to strengthen scientific practice that were implemented following the paper retractions.

More diversity can benefit science. The social diversity (e.g., gender and racial diversity) of teams can confer benefits to creativity, innovation, and accurate decision-making. Particularly important for the effort to improve science is that socially diverse scientific teams are more likely to generate innovation. Teams with more gender and cultural diversity are more likely to develop new products and introduce radical innovations to market (57, 58). In the domain of intellectual contributions, an analysis of the ethnic identities of authors of 1.2 million scientific papers published between 1985 and 2008 found that papers written by diverse groups received more citations and had higher impact factors than papers written by people from the same ethnic group (59). Yet the mere presence of socially diversity is not always sufficient to foster equal participation of diverse social groups. Indeed, the potential of social diversity often goes untapped, leading to null or negative results on group performance (60–63). For example, a large-scale analysis of contemporary scientific articles found that women were significantly more likely to be associated with technical tasks, whereas men were associated with conceptual tasks (64). Similarly, in gender-diverse engineering teams, female students were underrepresented in presenting technical content while male students were overrepresented (65).

To capitalize on the potential of social diversity, contexts need to directly address the challenges that can accompany social diversity. For example, interactions and communication within diverse teams may be more difficult (66–68). There is great potential of social diversity to produce better and more innovative outcomes, particularly in complex tasks: Socially diverse teams encode and process information more accurately (69), especially when sharing disparate facts is a requirement for success (70). The mere presence of people from socially diverse backgrounds alters the cognition and behavior of majority group members that lead to more improved and accurate thinking and communication (71). In the presence of social diversity, majority group members raise more facts and make fewer factual errors; when errors are made, they are more likely to be corrected (71). When questions and dissent are raised in socially diverse teams, it provokes more thought and consideration than when the exact same concerns are raised in homogenous teams (72). The presence of other underrepresented group members can foster greater participation from other underrepresented group members: One example is that gender-diverse teams with more women foster women's active participation in team projects, whereas teams that are comprised of mostly men often render women silent (67).

Lack of diversity also comes with costs: several high-profile failures underscore the need for social diversity on scientific teams. For example, with no women on engineering and development teams, heart valves and seat belts are made that only fit men's bodies (and significantly increase mortality rates for women) (73); voice recognition software that only recognizes the voices of men (73) and image recognition software that tags Black people as apes (74). Including and heeding the voices and experiences of a range of people can foster outcomes that benefit a wider range of people.

The emerging movements to improve science. The psychological and brain sciences (PBS) are at the forefront of efforts to redefine the rules and standards of science (75, 76). There is much to learn from this emerging movement and several other fields (77–81) are already taking similar stock, including biostatistics (82, 83), computer science (84), and medicine (85, 86). Consistent with a communal, prosocial cultural framework, fields with more established interdependent and collaborative scientific practices have enjoyed a positive evolution in scientific quality and discovery. For example, the movement toward team science can be observed in theoretical and experimental physics where investigative necessity has promoted large-scale consortia and successful models of scientific collaboration (87). With respect to structural biology, the establishment of standards for data sharing and deposition (see for example CCP4 and RCSB) coincided with broader participation of women in the field.

In sum, Open Science has the seed of a communal and sharing culture that, if cultivated, may be especially attractive to women and other underrepresented groups. We believe that the collaborative forward-looking aspect of open science has the potential to facilitate diversity and inclusiveness in two ways. First, the sharing of code, data and resources lowers the barriers and entry cost to the participation in science, thus establishing a more equal playing field and enhancing the inclusion of underrepresented groups - for example, scientists working in minority-serving institutions with less access to funding and other resources (88). Second, a culture of sharing, interdependence, and collaboration is consistent with research (cited above) that suggests these cultural features are more attractive to women, people of color, and communally-oriented men. In contrast, the culture that has come to characterize some of the reproducibility focus appears to be consistent with individualistic values that has traditionally turned away these underrepresented groups. An important speculation not fully addressed by our analyses is that women may have a reduced participation in the reproducibility movement because they might be less willing to adopt a finger-pointing posture (personal communication, M. McNutt). This is consistent with our claims on women choosing to participate in more open, communal cultures. More research will be necessary to address this issue, including overcoming the difficulties in ascertaining author gender (89).

Some aspects of the movements to improve science have explicitly focused on cultural values and practices to promote inclusivity. For example, the Society for the Improvement of Psychological Science explicitly includes working toward an inclusive culture in its mission statement, and the online methods and practices discussion group PsychMAP was founded to provide a more collaborative and communal space for discussion (see community ground rules). To be sure, reflecting and learning from *within* a cultural shift is difficult. The analysis we offer here suggests that we can still do more to improve science through social diversity. We propose that the benefits of team science will be realized when such teams are both socially and intellectually diverse and operate in contexts that welcome and pursue diversity, so that innovation, creativity, and the quality of science can flourish–despite an initial period of adjustment and discomfort. Science needs the participation of women and other underrepresented groups. The goals and ideals of open science have the potential to promote diversity and broader scientific participation. However, the promise of these emerging cultural trends is not yet a certainty; indeed, some features of the dominant scientific culture can deter participation among the very individuals who may contribute to the strength of diverse thinking. By fostering cultural change toward prosocial values, sharing, education, and cross-disciplinary cooperation, rather than independence and competitiveness, the movement to improve science may lead to greater knowledge generation, democratization, and inclusiveness in science.

Specific steps can and are being made to facilitate and advance the diversity we are promoting. Departments, institutions, and professional societies can create communal and prosocial structures for open science, such as open infrastructure and initiatives to allow for establishing educational networks, training, and resources and data sharing. Other specific examples include: Development of Transparency and Openness Promotion (TOP) Guidelines (16) and the establishment of cloud-based platforms and associated user communities for research asset sharing (see examples in PBS, OpenNeuro.org (90), brainlife.io (91, 92) and Open Science Framework (16)). Individual researchers can learn about the who, when, how and why of their teams, including attending to the range of people presented, identifying opportunities to include diverse voices, and analyzing reasons for and barriers or groups or individuals to participate in the team. Organizations that highlight the collaborative and communal aspects of scientific processes and success can feature connections in science, acknowledge how others help overcome stumbling blocks, and reward teams that embody the values of open science. Each researcher can work toward broadening collaboration and mentoring networks both in terms of collaborating with a range of established researchers and actively including underrepresented groups. We encourage readers and all members of the scientific community to embrace a learning mindset regarding team science and socially diverse teams: Science continually has more to teach, and the rewards of a cultural shift are not free; they come from investments of time, energy, and understanding.

METHODS

Data sources. A total of 11,338 original papers were collected using the snapshot of Microsoft Academic Graph (MAG; https://academic.microsoft.com) on Feb 23, 2018. To collect the datasets, we searched MAG for all publications with specific "field of study tags" as "open science" or "reproducibility." Among all the records, only 68 papers were categorized as both "open science" and "reproducibility". Moreover, of the 36,296 unique author ids represented in these literatures, very few (N = 457) have authored in both literatures. These findings suggest that the two literatures are developing rather independently. For the purposes of our analyses, we removed papers that were categorized as both "open science" and "reproducibility" to avoid double-counting papers and skewing analyses. Among the remaining records, we only considered formal published papers of the type "journal" or "conference". The resulting dataset includes 3,431 Open Science papers and 7,839 Reproducibility papers.

Among the remaining records, we only considered formal published papers of the type "journal" or "conference" (document types "book", "book chapter" and "patent" were removed). We also removed 43 papers with duplicate titles. We examined the remaining number of papers being published in each year within each literature, shown in the left panel of **Figure S1**. As very few Open Science papers were published prior to 2010, and few papers in either field have been published in 2018, we only use data for papers published between 2010 and 2017, which includes 879 Open Science papers and 2,047 Reproducibility papers. This is the final dataset used for all analyses, except where otherwise noted in the text.

Data compiled for the analyses can be found at https://osf.io/97vcx and code used for this work is available at <GITHUB-URL>.

Based on the sample between 2010 and 2017, we constructed the paper co-authorship networks for 879 Open Science papers and 2,047 Reproducibility papers. Each node represents a scientific article. Two nodes share an edge if at least one author appears in both papers. Based on MAG author ids, we identified 3,157 unique author names in the open science literature and 8,766 in the reproducibility literature. In the open science literature, the network contains 389 edges (i.e., pairs of papers with at least one author in common), and 856 edges in the reproducibility literature.

Network analysis. For both networks, we measured the following quantities:
(1) Edge density. For an undirected network with n nodes and m edges, the edge density is defined as:

$$\rho = \frac{m}{(n \times (n-1))/2}.$$

To test whether the open science network has higher edge density than the reproducibility network, we conducted a one-sided Fisher's exact test. We assumed a binomial edge generation process between all pairs of nodes and tested the hypothesis that the odds ratio of the two networks is greater than one. We estimated the odds ratio using the edge density of both networks,

$$\frac{\rho_1(1-\rho_2)}{\rho_2(1-\rho_1)},$$

Where ρ_1 represents the edge density of the open science network and ρ_2 for the reproducibility network. The odds ratio test was used to handle the small values of the network density (0.057% and 0.047%), opposed to a test utilizing a linear scale. The test rejects the null hypothesis that the open science network does *not* have higher edge density than the reproducibility network with a p-value of 7.35e-05.

(2) Connected components. We performed an additional analysis to estimate how connected (or isolated) the subcomponents of each network are. For an undirected network, a connected component is defined as a maximal subgraph in which any two nodes are connected to each other by a sequence of edges. In our case, both networks are sparse with many separate connected components. We compared the two networks in terms of the size of the largest connected component, as well as the average component size, which is defined as the network size divided by the number of connected components. The connected components analysis is conducted using the software Gephi (29).

Gender representation analysis (Figures 2, 3 and 4). We performed a traditional gender (male, female) analysis by identifying the gender of the first and last authors given their name. To do so, we used the *gender* R package (https://github.com/ropensci/gender; (93) to determine the probability of the first and last author to be a female. The *gender* package uses historical data on gender to predict the gender of a person based on their given name(s) and birth year or year range. For each paper, we assumed birth year to be such that the author would be between the ages of 25 and 65 at the time of publication. To identify the first name of each author, we first identified the component of each author name by assuming that each name component was separated by one space in the data. We then considered the first and middle names (when available) and excluded all other initials to perform gender detection. We computed the probability of being female for each author with at least one full (non-initial) first or middle name part. Authors with probability over 0.5 were labeled “female” and those with probability below 0.5 were labeled “male.” We used the "ssa" option of the gender package, which looks up names based from the U.S. Social Security Administration baby name data from the period 1932 to 2012.

For **Figures 3** and **4**, we labeled papers as having a female in a high-status author position if either the first or last author was labeled “female” using the method described above. We excluded papers with unknown high-status female authorship, which includes papers with both first and last author labeled “unknown” and papers with one position “male” and the other “unknown.” We excluded single-author papers, since a lower proportion of those would be expected to have female high-status authorship compared with multi-author papers (since in a probabilistic sense there are two “chances” to achieve high-status authorship in multi-author papers but only one “chance” in single-author papers). In **Figure 3**, we also excluded papers with more than 15 authors for the sake of visualization.

For **Figure 4**, we performed logistic regression analysis to quantify how the rates of female high-status authorship in multi-author papers varied over time and by team size within each literature. We included a spline term for team size within each literature, given the evidence for a non-linear relationship between team size and rates of female lead authorship. We excluded 28 Open Science papers and 40 Reproducibility papers with more than 12 authors to avoid undue influence on the estimation of these spline terms. The resulting dataset consisted of 454 Open Science papers and 955 Reproducibility papers.

Specifically, we fit a logistic regression model relating the log-odds of having a female in a high-status author position to the year of publication, the number of authors (using a flexible spline term), the type of publication (conference proceedings or journal article) and the literature each paper belongs to. We allowed the effects of year of publication and number of authors to be determined separately for each literature through interaction terms. We estimated the model coefficients using the R gam function from the mgcv package using a binomial family with logit link. This function represents smooth coefficient curves as penalized splines and uses generalized cross-validation to estimate the smoothness of each curve (94). Specifically, we fit the model

$$\log\left\{\frac{Pr(Y_i=1)}{1-Pr(Y_i=1)}\right\} = \beta_0 + \beta_1 Rep_i + \beta_2 Year_i + \beta_3 Year_i Rep_i + f_1(Authors_i) + f_2(Authors_i Rep_i) + \beta_4 Conf_i + \varepsilon_i$$
$$\varepsilon_i \sim N(0, \sigma^2)\,,$$

where $Y_i = 1$ if paper i has a female in a high-status author position, $Rep_i = 1$ if paper i belongs to the Reproducibility literature, $Year_i$ is the year of publication (centered at 2017), $Authors_i$ is the team size (centered at 2, the minimum value for multi-author papers), and $Conf_i = 1$ if the paper is a conference proceeding. The functions $f_1()$ and $f_2()$ are smooth coefficient curves that map team size to the log-odds of having a female high-status author in each literature, given fixed values of the other coefficients.

Based on the estimated regression coefficients and standard errors, we estimated the log-odds of having a female in a lead authorship position given specific sets of predictor variables, along with Normal 95% confidence intervals (CIs). We then transformed the log-odds and CIs to odds and probabilities for better interpretability. **Supplementary Table S1** reports estimates and confidence intervals on the odds scale for each parametric (i.e., non-spline) coefficient. In short, we find that the effect of belonging to the reproducibility literature is negative with an estimate of 0.393, representing approximately 61% reduced odds of having a woman in a high-status position compared with papers in open science for a given team size, year of publication and manuscript type. The effect of later publication year is positive for open science papers but negative for reproducibility papers. All parametric coefficients are statistically significant at the 0.05 level.

The effects of publication year and team size are explored in further detail by examining the predicted probabilities of having a woman in a high-status position as year and team size vary. **Figure 4** in the main text displays the estimates and 95% confidence intervals for the probability of having a female in a high-status position for different values of these variables. Confidence intervals on the probability scale are constructed by applying the inverse logit transformation (i.e., $p(x) = exp\{x/(1-x)\}$) to the Normal 95% confidence interval on the logit scale. The estimates and confidence intervals therefore represent predicted probabilities. In the **left panel of Figure 4**, we fix Conference = FALSE and Year = 2017, while allowing team size to vary for each literature. The results show that for open science papers there is a negative effect of team size, with smaller teams having slightly higher probability of having a female in a high-status position. For reproducibility papers, there is a nonlinear effect of team size, with the probability of having a female in a high-status position being markedly lower for small teams and peaking for teams with approximately 7 authors before declining slightly. In the **right panel of Figure 4**, we fix Conference = FALSE and Team Size = 4 (near the mean value), while allowing year of publication to vary for each literature. We observe a striking difference in the effect of year of publication for open science and reproducibility papers, with an increasing trend over time for open science papers and a slightly decreasing trend over time for reproducibility papers. This suggests increasing participation of women in high-status positions within the open science literature and a decline or stagnation in the reproducibility literature.

Finally, we repeated the regression analysis controlling for academic fields of study. Since each paper lists multiple fields and there are over 2,000 unique fields listed, we considered the 30 most frequently appearing fields, which constitute over 44% of all field appearances. Of these 30, we removed several fields not considered traditional fields of study ("publishing", "workflow", "data sharing", "reproducibility", "open science", "repeatability", "open data", "data mining", "intraclass correlation") and combined several medicine-related fields into a single “medicine” field ("anesthesiology", "cardiology", "diabetes mellitus", "internal medicine", "surgery", "alternative medicine", "physical therapy", "pathology", "radiology", "medicine"). This resulted in 12 academic fields of study ("medicine", "artificial intelligence", "management science", "analytical chemistry", "bioinformatics", "engineering", "knowledge management", "psychology", "software", "biology", "statistics", "computer science"). For each field, we constructed an indicator variable equalling 1 if the field was listed by each paper among its fields of study and 0 otherwise. We controlled for field of study by including these indicator variables as binary covariates in the model specified above. Of the 1,409 papers included in the original model, only 98 did not list any of these fields and were excluded from this model.

Supplementary Table S2 reports coefficient estimates and 95% confidence intervals on the odds scale for each parametric (non-spline coefficient). In short, the effect of belonging to the Reproducibility literature remains negative and statistically significant, and the trend over time is again positive within the Open Science literature and slightly negative within the Reproducibility literature.

Text Analyses of Abstracts (Figure 5). Starting with the 2,926 papers from both open science and reproducibility described above, we first removed papers without available abstracts (205 Open Science and 815 Reproducibility papers), then removed those with non-English titles (79 Open Science and 63 Reproducibility papers) as determined using the R textcat package (95). The resulting dataset used in the text analysis consisted of 1,764 papers, including 595 Open Science papers and 1,169 Reproducibility papers. We then performed standard text preprocessing and removed stopwords, stemming, punctuations, and converting the text to lowercase using the SentimentAnalysis R package. We measured prosocial constructs in the text by counting the frequency of occurrence of 127 words in a validated dictionary (95) (e.g., contribute, encourage, help, nurture, see **Supplementary Table S3**). This dictionary has been shown to have high agreement with human judges (r = 0.67) (96). The prosocial word density is calculated as the ratio of the number of prosocial words over total number of words in each abstract.

REFERENCES

1. The Team Science Revolution (2014) *Science* | *AAAS*. Available at: http://www.sciencemag.org/careers/2011/06/team-science-revolution [Accessed February 28, 2018].

2. Markus HR, Kitayama S (2010) Cultures and Selves: A Cycle of Mutual Constitution. *Perspect Psychol Sci* 5(4):420–430.

3. Luthar SS (2017) Doing for the Greater Good: What Price, in Academe? *Perspect Psychol Sci* 12(6):1153–1158.

4. Mitchneck B, Smith JL, Latimer M (2016) A recipe for change: Creating a more inclusive academy. *Science* 352(6282):148–149.

5. Syed M (2017) Why Traditional Metrics May Not Adequately Represent Ethnic Minority Psychology. *Perspect Psychol Sci* 12(6):1162–1165.

6. Roediger HL 3rd (2016) Varieties of Fame in Psychology. *Perspect Psychol Sci* 11(6):882–887.

7. Diekman AB, Brown ER, Johnston AM, Clark EK (2010) Seeking congruity between goals and roles: a new look at why women opt out of science, technology, engineering, and mathematics careers. *Psychol Sci* 21(8):1051–1057.

8. Feist GJ (2016) Intrinsic and Extrinsic Science: A Dialectic of Scientific Fame. *Perspect Psychol Sci* 11(6):893–898.

9. Zárate MA, Hall GN, Plaut VC (2017) Researchers of Color, Fame, and Impact. *Perspect Psychol Sci* 12(6):1176–1178.

10. Diekman AB, Clark EK, Johnston AM, Brown ER, Steinberg M (2011) Malleability in communal goals and beliefs influences attraction to stem careers: evidence for a goal congruity perspective. *J Pers Soc Psychol* 101(5):902–918.

11. Eagly AH, Karau SJ (2002) Role congruity theory of prejudice toward female leaders. *Psychol Rev* 109(3):573–598.

12. Murphy MC, Steele CM, Gross JJ (2007) Signaling threat: how situational cues affect women in math, science, and engineering settings. *Psychol Sci* 18(10):879–885.

13. Baker M (2015) Over half of psychology studies fail reproducibility test. *Nature News* 27.

14. Spies JR The Open Science Framework: Improving Science by Making It Open and Accessible. doi:10.18130/v35c6x.

15. Poldrack RA, et al. (2013) Toward open sharing of task-based fMRI data: the OpenfMRI project. *Front Neuroinform* 7:12.

16. Nosek BA, et al. (2015) Promoting an open research culture. *Science* 348(6242):1422–1425.

17. Avesani P, et al. (2018) The open diffusion data derivatives, brain data upcycling via integrated publishing of derivatives and reproducible open cloud services. doi:10.31234/osf.io/y82n7.

18. Staff W (2017) Diversity in Open Source Is Even Worse Than in Tech Overall. *WIRED*. Available at: https://www.wired.com/2017/06/diversity-open-source-even-worse-tech-overall/ [Accessed February 26, 2018].

19. Nosek B Diversity and inclusion in the open science movement. *Center for Open Science*. Available at: https://cos.io/blog/how-can-we-improve-diversity-and-inclusion-open-science-movement/ [Accessed February 26, 2018].

20. Cech EA, Metz A, Smith JL, deVries K (2017) Epistemological Dominance and Social Inequality. *Sci Technol Human Values* 42(5):743–774.

21. Schneider MC, Holman MR, Diekman AB, McAndrew T (2016) Power, Conflict, and Community: How Gendered Views of Political Power Influence Women's Political Ambition. *Polit Psychol* 37(4):515–531.

22. Thoman DB, Brown ER, Mason AZ, Harmsen AG, Smith JL (2015) The Role of Altruistic Values in Motivating Underrepresented Minority Students for Biomedicine. *Bioscience* 65(2):183–188.

23. Valantine HA, Collins FS (2015) National Institutes of Health addresses the science of diversity. *Proc Natl Acad Sci U S A* 112(40):12240–12242.

24. Intemann K (2009) Why Diversity Matters: Understanding and Applying the Diversity Component of the National Science Foundation's Broader Impacts Criterion. *Social Epistemology* 23(3-4):249–266.

25. Page SE (2008) *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies* (Princeton University Press) Available at: https://market.android.com/details?id=book-hJRu4O8q1xwC.

26. Sinha A, et al. (2015) An Overview of Microsoft Academic Service (MAS) and Applications. *Proceedings of the 24th International Conference on World Wide Web*, WWW '15 Companion. (ACM, New York, NY, USA), pp 243–246.

27. Newman M (2018) *Networks* (Oxford University Press) Available at: https://market.android.com/details?id=book-YdZjDwAAQBAJ.

28. Barabási A-L, Pósfai M (2016) *Network Science* (Cambridge University Press) Available at: https://market.android.com/details?id=book-iLtGDQAAQBAJ.

29. Bastian M, Heymann S, Jacomy M, Others (2009) Gephi: an open source software for exploring and manipulating networks. *Icwsm* 8(2009):361–362.

30. Frimer JA, Aquino K, Gebauer JE, Zhu L (lei), Oakes H (2015) A decline in prosocial language helps explain public disapproval of the US Congress. *Proc Natl Acad Sci U S A* 112(21):6591–6594.

31. Bian L, Leslie S-J, Murphy MC, Cimpian A (2018) Messages about brilliance undermine women's interest in educational and professional opportunities. *J Exp Soc Psychol*. doi:10.1016/j.jesp.2017.11.006.

32. Leslie S-J, Cimpian A, Meyer M, Freeland E (2015) Expectations of brilliance underlie gender distributions across academic disciplines. *Science* 347(6219):262–265.

33. Rattan A, et al. (2017) Meta-lay Theories of Scientific Potential Drive Underrepresented Students' Sense of Belonging to Science, Technology, Engineering, and Mathematics (STEM). Available at: https://www.krishnasavani.com/wordpress/wp-content/uploads/2017/12/Rattan_Savani_et_al_2018.pdf.

34. AAMCNews (2016) Available at: https://news.aamc.org/ [Accessed February 27, 2018].

35. McGovern V (2009) Perspective: How to Succeed in Big Science and Still Get Tenure. *Science*. doi:10.1126/science.caredit.a0900092.

36. Halpern DF (2017) Whither Psychology. *Perspect Psychol Sci* 12(4):665–668.

37. Smith SW (1987) IRIS—A university consortium for seismology. *Rev Geophys* 25(6):1203.

38. Larivière V, Gingras Y, Sugimoto CR, Tsou A (2014) Team size matters: Collaboration and scientific impact since 1900. *Journal of the Association for Information Science and Technology* 66(7):1323–1332.

39. McPherson M, Smith-Lovin L, Cook JM (2001) Birds of a Feather: Homophily in Social Networks. *Annu Rev Sociol* 27(1):415–444.

40. Adams J (2013) Collaborations: The fourth age of research. *Nature* 497(7451):557–560.

41. Ginther DK, et al. (2011) Race, ethnicity, and NIH research awards. *Science* 333(6045):1015–1019.

42. Wuchty S, Jones BF, Uzzi B (2007) The increasing dominance of teams in production of knowledge. *Science* 316(5827):1036–1039.

43. Jones BF, Wuchty S, Uzzi B (2008) Multi-university research teams: shifting impact, geography, and stratification in science. *Science* 322(5905):1259–1262.

44. Disis ML, Slattery JT (2010) The road we must take: multidisciplinary team science. *Sci Transl Med* 2(22):22cm9.

45. Post C, et al. (2009) Capitalizing on Thought Diversity for Innovation. *Research-Technology Management* 52(6):14–25.

46. Singh J, Fleming L (2010) Lone Inventors as Sources of Breakthroughs: Myth or Reality? *Manage Sci* 56(1):41–56.

47. Fox CW, Timothy Paine CE, Sauterey B (2016) Citations increase with manuscript length, author number, and references cited in ecology journals. *Ecol Evol* 6(21):7717–7726.

48. Gazni A, Didegah F (2011) Investigating different types of research collaboration and citation impact: a case study of Harvard University's publications. *Scientometrics* 87(2):251–265.

49. Katz JS, Martin BR (1997) What is research collaboration? *Res Policy* 26(1):1–18.

50. Azoulay P (2019) Small research teams "disrupt" science more radically than large ones. *Nature*. doi:10.1038/d41586-019-00350-3.

51. Wu L, Wang D, Evans JA (2019) Large teams develop and small teams disrupt science and technology. *Nature*. doi:10.1038/s41586-019-0941-9.

52. Ferry G (2014) History: Women in crystallography. *Nature* 505(7485):609–611.

53. Borrell B (2009) Fraud rocks protein community. *Nature* 462(7276):970.

54. Read RJ, et al. (2011) A new generation of crystallographic validation tools for the protein data bank. *Structure* 19(10):1395–1412.

55. Francl M (2014) Seeding crystallography. *Nat Chem* 6(10):842–844.

56. Sheltzer JM, Smith JC (2014) Elite male faculty in the life sciences employ fewer women. *Proc Natl Acad Sci U S A* 111(28):10107–10112.

57. Díaz-García C, González-Moreno A, Jose Sáez-Martínez F (2013) Gender diversity within R&D

teams: Its impact on radicalness of innovation. *Innovations* 15(2):149–160.

58. Nathan M, Lee N (2013) Cultural Diversity, Innovation, and Entrepreneurship: Firm-level Evidence from London. *Econ Geogr* 89(4):367–394.

59. Freeman RB, Huang W (2015) Collaborating with People Like Me: Ethnic Coauthorship within the United States. *J Labor Econ* 33(S1):S289–S318.

60. Galinsky AD, et al. (2015) Maximizing the Gains and Minimizing the Pains of Diversity: A Policy Perspective. *Perspect Psychol Sci* 10(6):742–748.

61. van Dijk H, van Engen ML, van Knippenberg D (2012) Defying conventional wisdom: A meta-analytical examination of the differences between demographic and job-related diversity relationships with performance. *Organ Behav Hum Decis Process* 119(1):38–53.

62. Eagly AH (2016) When Passionate Advocates Meet Research on Diversity, Does the Honest Broker Stand a Chance?: Passionate Advocates and Diversity Research. *J Soc Issues* 72(1):199–222.

63. Apfelbaum EP, Phillips KW, Richeson JA (2014) Rethinking the Baseline in Diversity Research: Should We Be Explaining the Effects of Homogeneity? *Perspect Psychol Sci* 9(3):235–244.

64. Macaluso B, Larivière V, Sugimoto T, Sugimoto CR (2016) Is Science Built on the Shoulders of Women? A Study of Gender Differences in Contributorship. *Acad Med* 91(8):1136–1142.

65. Meadows LA, Sekaquaptewa D (2013) The influence of gender stereotypes on role adoption in student teams. *Proc. 120th ASEE Annual Conf. Exposition* (American Society for Engineering Education Washington, DC), pp 1–16.

66. Rink FA, Ellemers N (2009) Temporary versus permanent group membership: how the future prospects of newcomers affect newcomer acceptance and newcomer influence. *Pers Soc Psychol Bull* 35(6):764–775.

67. Dasgupta N, Scircle MM, Hunsinger M (2015) Female peers in small work groups enhance women's motivation, verbal participation, and career aspirations in engineering. *Proc Natl Acad Sci U S A* 112(16):4988–4993.

68. van Knippenberg D, De Dreu CKW, Homan AC (2004) Work group diversity and group performance: an integrative model and research agenda. *J Appl Psychol* 89(6):1008–1022.

69. Levine SS, et al. (2014) Ethnic diversity deflates price bubbles. *Proc Natl Acad Sci U S A* 111(52):18524–18529.

70. Phillips KW, Liljenquist KA, Neale MA (2009) Is the pain worth the gain? The advantages and liabilities of agreeing with socially distinct newcomers. *Pers Soc Psychol Bull* 35(3):336–350.

71. Sommers SR (2006) On racial diversity and group decision making: identifying multiple effects of racial composition on jury deliberations. *J Pers Soc Psychol* 90(4):597–612.

72. Antonio AL, et al. (2004) Effects of racial diversity on complex thinking in college students. *Psychol Sci* 15(8):507–510.

73. Science Case Studies | Gendered Innovations Available at: https://genderedinnovations.stanford.edu/innovations-in-design/cases.html [Accessed February 28, 2018].

74. Lee W (2015) How tech's lack of diversity leads to racist software. *SFGate*. Available at: https://www.sfgate.com/business/article/How-tech-s-lack-of-diversity-leads-to-racist-6398224.php

[Accessed February 28, 2018].

75. Ledgerwood A (2014) Introduction to the Special Section on Advancing Our Methods and Practices. *Perspect Psychol Sci* 9(3):275–277.

76. Psychology's Credibility Crisis | DiscoverMagazine.com *Discover Magazine*. Available at: http://discovermagazine.com/2014/jan-feb/46-research-report-repeat-with-brian-nosek [Accessed February 28, 2018].

77. Baker M (2016) 1,500 scientists lift the lid on reproducibility. *Nature* 533(7604):452–454.

78. Lowndes JSS, et al. (2017) Our path to better science in less time using open data science tools. *Nat Ecol Evol* 1(6):160.

79. Reilly D, et al. (1994) Is evidence for homoeopathy reproducible? *Lancet* 344(8937):1601–1606.

80. Baker M, Dolgin E (2017) Cancer reproducibility project releases first results. *Nature* 541(7637):269–270.

81. Mullard A (2017) Cancer reproducibility project yields first results. *Nat Rev Drug Discov* 16(2):77.

82. Buckheit JB, Donoho DL (1995) WaveLab and Reproducible Research. *Wavelets and Statistics*, Lecture Notes in Statistics. (Springer, New York, NY), pp 55–81.

83. Buck S (2015) Solving reproducibility. *Science* 348(6242):1403.

84. Missing data hinder replication of artificial intelligence studies (2018) *Science* | *AAAS*. doi:10.1126/science.aat3298.

85. Arturo Casadevall FCF (2010) Reproducible Science. *Infect Immun* 78(12):4972.

86. Engber D (2016) Think Psychology's Replication Crisis Is Bad? Welcome to the One in Medicine. *Slate Magazine*. Available at: http://www.slate.com/articles/health_and_science/future_tense/2016/04/biomedicine_facing_a_worse_replication_crisis_than_the_one_plaguing_psychology.html [Accessed February 28, 2018].

87. Abbott BP, et al. (2016) Observation of Gravitational Waves from a Binary Black Hole Merger. *Phys Rev Lett* 116(6):061102.

88. Matthews CM (1993) Federal research and development funding at Historically Black Colleges and Universities (Congressional Research Service, Library of Congress). Available at: http://research.policyarchive.org/18821.pdf.

89. Sugimoto CR, Ahn Y-Y, Smith E, Macaluso B, Larivière V (2019) Factors affecting sex-related reporting in medical research: a cross-disciplinary bibliometric analysis. *Lancet* 393(10171):550–559.

90. Gorgolewski K, Esteban O, Schaefer G, Wandell B, Poldrack R (2017) OpenNeuro—a free online platform for sharing and analysis of neuroimaging data. *Organization for Human Brain Mapping Vancouver, Canada* 1677. Available at: https://f1000research.com/assets/download/1114354.

91. Pestilli F (2018) Human white matter and knowledge representation. *PLoS Biol* 16(4):e2005758.

92. Hayashi S, Avesani P, Pestilli F (2017) *Open Diffusion Data Derivatives* (brainlife.io) doi:10.25663/BL.P.3.

93. Blevins C, Mullen L (2015) Jane, John... Leslie? A Historical Method for Algorithmic Gender Prediction. *DHQ: Digital Humanities Quarterly* 9(3). Available at:

http://search.ebscohost.com/login.aspx?direct=true&profile=ehost&scope=site&authtype=crawler&jrnl=19384122&AN=125676755&h=ypzTlFHkz6dlcPoQQ%2BwyWw3D4XZUuS0fiG95SaztiSOC3iwKyTeRzpmVXBlk7eyM6W3DgXxswEPUfBiwCnpojw%3D%3D&crl=c.

94. Wood SN, Pya N, Säfken B (2016) Smoothing Parameter and Model Selection for General Smooth Models. *J Am Stat Assoc* 111(516):1548–1563.

95. Hornik K, et al. (2013) The textcat Package for n-Gram Based Text Categorization in R. *Journal of Statistical Software, Articles* 52(6):1–17.

96. Frimer JA, Schaefer NK, Oakes H (2014) Moral actor, selfish agent. *J Pers Soc Psychol* 106(5):790–802.