



Open science, communal culture, and women's participation in the movement to improve science

Mary C. Murphy^{a,1,2}, Amanda F. Mejia^{b,1}, Jorge Mejia^{c,1}, Xiaoran Yan^{d,1}, Sapna Cheryan^e, Nilanjana Dasgupta^f, Mesmin Destin^{g,h,i}, Stephanie A. Fryberg^j, Julie A. Garcia^k, Elizabeth L. Haines^l, Judith M. Harackiewicz^m, Alison Ledgerwoodⁿ, Corinne A. Moss-Racusin^o, Lora E. Park^p, Sylvia P. Perry^{g,h,q}, Kate A. Ratliff^r, Aneeta Rattan^s, Diana T. Sanchez^t, Krishna Savani^u, Denise Sekaquaptewa^j, Jessi L. Smith^{v,w}, Valerie Jones Taylor^{x,y}, Dustin B. Thoman^z, Daryl A. Wout^{aa}, Patricia L. Mabry^{bb,3}, Susanne Ressler^{cc,dd,3}, Amanda B. Diekmann^{a,3}, and Franco Pestilli^{ae,3}

^aDepartment of Psychological and Brain Sciences, Indiana University Bloomington, Bloomington, IN 47405; ^bDepartment of Statistics, Indiana University Bloomington, Bloomington, IN 47408; ^cKelley School of Business, Indiana University Bloomington, Bloomington, IN 47405; ^dNetwork Science Institute, Indiana University Bloomington, Bloomington, IN 47408; ^eDepartment of Psychology, University of Washington, Seattle, WA 98195; ^fDepartment of Psychological and Brain Sciences, University of Massachusetts Amherst, Amherst, MA 01003; ^gDepartment of Psychology, Northwestern University, Evanston, IL 60208; ^hInstitute for Policy Research, Northwestern University, Evanston, IL 60208; ⁱSchool of Education & Social Policy, Northwestern University, Evanston, IL 60208; ^jDepartment of Psychology, University of Michigan–Ann Arbor, Ann Arbor, MI 48109; ^kDepartment of Psychology and Child Development, California Polytechnic State University, San Luis Obispo, CA 93407; ^lDepartment of Psychology, William Paterson University, Wayne, NJ 07470; ^mDepartment of Psychology, University of Wisconsin–Madison, Madison, WI 53706; ⁿDepartment of Psychology, University of California, Davis, CA 95616; ^oDepartment of Psychology, Skidmore College, Saratoga Springs, NY 12866; ^pDepartment of Psychology, University at Buffalo, The State University of New York, Buffalo, NY 14260; ^qDepartment of Medical Social Sciences, Northeastern University, Evanston, IL 60208; ^rDepartment of Psychology, University of Florida, Gainesville, FL 32611; ^sOrganisational Behaviour, London Business School, London NW1 4SA, United Kingdom; ^tDepartment of Psychology, Rutgers, The State University of New Jersey, Piscataway, NJ 08854; ^uLeadership, Management & Organisation, Nanyang Technological University, Singapore 639798; ^vOffice of Research, University of Colorado Colorado Springs, Colorado Springs, CO 80918; ^wDepartment of Psychology, University of Colorado Colorado Springs, Colorado Springs, CO 80918; ^xDepartment of Psychology, Lehigh University, Bethlehem, PA 18015; ^yAfricana Studies, Lehigh University, Bethlehem, PA 18015; ^zDepartment of Psychology, San Diego State University, San Diego, CA 92182; ^{aa}Department of Psychology, John Jay College of Criminal Justice, City University of New York, New York, NY 10019; ^{bb}Research Division, HealthPartners Institute, Bloomington, MN 55425; ^{cc}Department of Molecular and Cellular Biochemistry, Indiana University Bloomington, Bloomington, IN 47405; ^{dd}Department of Neuroscience, The University of Texas at Austin, Austin, TX 78712; and ^{ee}Department of Psychology, The University of Texas at Austin, Austin, TX 78712

Edited by Susan T. Fiske, Princeton University, Princeton, NJ, and approved July 27, 2020 (received for review December 7, 2019)

Science is undergoing rapid change with the movement to improve science focused largely on reproducibility/replicability and open science practices. This moment of change—in which science turns inward to examine its methods and practices—provides an opportunity to address its historic lack of diversity and noninclusive culture. Through network modeling and semantic analysis, we provide an initial exploration of the structure, cultural frames, and women's participation in the open science and reproducibility literatures ($n = 2,926$ articles and conference proceedings). Network analyses suggest that the open science and reproducibility literatures are emerging relatively independently of each other, sharing few common papers or authors. We next examine whether the literatures differentially incorporate collaborative, prosocial ideals that are known to engage members of underrepresented groups more than independent, winner-takes-all approaches. We find that open science has a more connected, collaborative structure than does reproducibility. Semantic analyses of paper abstracts reveal that these literatures have adopted different cultural frames: open science includes more explicitly communal and prosocial language than does reproducibility. Finally, consistent with literature suggesting the diversity benefits of communal and prosocial purposes, we find that women publish more frequently in high-status author positions (first or last) within open science (vs. reproducibility). Furthermore, this finding is further patterned by team size and time. Women are more represented in larger teams within reproducibility, and women's participation is increasing in open science over time and decreasing in reproducibility. We conclude with actionable suggestions for cultivating a more prosocial and diverse culture of science.

through greater rigor and transparency (e.g., open sharing of data, code, resources; standardized statistical procedures; preregistration). As with any revolution, a time of unrest can also be a time of opportunity. Indeed, researchers involved in the efforts to improve

Significance

Science is rapidly changing with the current movement to improve science focused largely on reproducibility/replicability and open science practices. Through network modeling and semantic analysis, this article provides an initial exploration of the structure, cultural frames of collaboration and prosociality, and representation of women in the open science and reproducibility literatures. Network analyses reveal that the open science and reproducibility literatures are emerging relatively independently with few common papers or authors. Open science has a more collaborative structure and includes more explicit language reflecting communality and prosociality than does reproducibility. Finally, women publish more frequently in high-status author positions within open science compared with reproducibility. Implications for cultivating a diverse, collaborative culture of science are discussed.

Author contributions: M.C.M., A.F.M., J.M., X.Y., P.L.M., S.R., A.B.D., and F.P. designed research; M.C.M., A.F.M., J.M., X.Y., P.L.M., S.R., A.B.D., and F.P. performed research; A.F.M., J.M., and X.Y. analyzed data; M.C.M., A.F.M., J.M., X.Y., S.C., N.D., M.D., S.A.F., J.A.G., E.L.H., J.M.H., A.L., C.A.M.-R., L.E.P., S.P.P., K.A.R., A.R., D.T.S., K.S., D.S., J.L.S., V.J.T., D.B.T., D.A.W., P.L.M., S.R., A.B.D., and F.P. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

This open access article is distributed under [Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 \(CC BY-NC-ND\)](https://creativecommons.org/licenses/by-nc-nd/4.0/).

¹M.C.M., A.F.M., J.M., and X.Y. contributed equally to this work.

²To whom correspondence may be addressed. Email: mcpsych@indiana.edu.

³P.L.M., S.R., A.B.D., and F.P. contributed equally to this work.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1921320117/-DCSupplemental>.

First published September 14, 2020.

open science | reproducibility | replicability | women | culture

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 primary goals: 1) understanding the flaws, weaknesses, and reproducibility of past scientific processes and findings (e.g., evaluating the strength of the evidence) and 2) improving research practices

science have acknowledged a gender diversity problem (2, 3), and this time of reform offers the opportunity to reinvent scientific culture in a more inclusive mode. If the movement to improve science perpetuates the traditional scientific culture that prioritizes independent, dominant, or adversarial values, it risks continuing to leave many talented individuals at the margins, feeling unwelcome and excluded (4)—exacerbating a global problem that the sciences are trying to solve (5–8). In its efforts to improve its methods and replicability, we wondered whether science might also be achieving improvements in the gender representation and inclusivity of the movement itself. This article applies cultural and network analysis to examine the emerging cultures in the movement to improve science—specifically in the reproducibility and open science literatures—and to investigate the representation of women in these emerging subcultures. We discuss implications of these different cultural avenues for science going forward.

In cultural analyses, the actions and cognitions of individuals both rise from and produce the norms and practices of groups and institutions (9). Further, the “who” and the “how” of cultural practices are inextricably intertwined: “how” a subculture operates influences “who” engages in the subculture, and “who” engages in the subculture influences “how” a subculture operates. The cultural practices of the current scientific reform movements influence who engages. The emerging reform movements 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 (10–12). The culture of science has long valued individual brilliance, competition, and a winner-take-all model of success (13). 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 (14).

The scientific practice of rewarding individual achievement has perhaps unwittingly fostered a more independent, competitive culture that ignores and possibly even disincentivizes cooperation (15, 16). These cultural practices have implications for who joins and advances within scientific fields. For example, the perceived lack of prosocial and collaborative culture in STEM has been shown to deter women especially (14, 17). 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 critiqued 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 in the numerical minority) against established figures, because this conflict-prone stance violates gender role expectations (18). Women who are perceived as self-promoting or aggressive face more negative evaluations than their male counterparts (19); thus, engaging in critiques or debates can elicit more backlash toward women than men, and the mere anticipation of backlash can inhibit women’s engagement in these spheres. Second, women may prefer a collective approach for pragmatic and principled reasons. Pragmatically, there is psychological safety in numbers (20–22), and women’s critiques may be more likely to be offered and listened to when they are part of a larger scientific team. Further, because combative and adversarial behaviors are perceived as masculine, women may be less socialized to engage in these behaviors than men and/or view them as off-putting and less likely to be productive (23). In principle, a collectivist orientation may disfavor challenges to the establishment when framed as for the benefit of the challenger (i.e., gaining recognition) rather than for the collective good (i.e., improving and advancing science).

However, we draw attention to another causal pathway as well: subcultures that include a larger proportion of women (or other underrepresented group members) could engage in different practices than more homogenous subcultures. For example,

legislative bodies that include greater proportions of women legislators engage more with policies related to education and health care (24–26). Culture is a cyclical process, and thus greater inclusion and advancement of women foster norms and behaviors that in turn can contribute to increasing gender diversity (6, 27, 28).

The movement to improve science, to date, can be characterized by two contrasting motifs—both aimed to improve science. One focus centers on the assessment of the reproducibility and replicability of previously published scientific results. We note that the National Academies of Sciences, Engineering, and Medicine has only recently formalized a distinction between reproducibility and replicability (29). Before this formalization, the two terms had historically been used with different conventions in different fields, with a prevalence of the term reproducibility (29–36). For this reason, our analysis (that uses historical data across fields) does not separate the two; instead, throughout the report, we use the term “reproducibility” to refer to the literature that we analyze.* A second approach aimed to improve science consists of “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 (37–39). For shorthand, we refer to these two literatures as “reproducibility” and “open science.” Indeed, both literatures aim to improve science, are led by scientists, and engage in deep analysis and critique of current scientific practices while offering guidance and suggestions for how to improve scientific practices. Here, we explored whether the reproducibility and open science literatures exhibit different 1) collaborative structures, 2) explicitly prosocial foci, and 3) engagement of female scientists. We anticipated that this initial investigation would reveal evidence of different emerging cultures in the reproducibility vs. open science literatures—with implications for the future representation and practices of these movements.

Our team conducted network analyses of the open science and reproducibility literatures and found that these literatures have few common papers and authors—suggesting these improvement approaches have developed relatively independently from each other. Given this, we compared these literatures for hallmarks of collaborative and prosocial culture. We find a more interconnected authorship network within open science compared with reproducibility, and semantic text analyses of article abstracts reveal 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 prosociality compared with reproducibility. We then examine the configuration of women’s participation in these literatures. We find patterns of women’s participation consistent with the theoretical idea that women’s participation is less constrained in more collaborative and prosocial cultures (i.e., in open science than in reproducibility). Women scholars are more likely to occupy high-status author positions (taking the first or last author position) within open science compared with reproducibility (see Fig. 3); further, women’s high-status authorship occurs less frequently in smaller teams within reproducibility (compared with open science). In larger teams—that might offer greater collective safety or communal purpose—

*Today, it is acknowledged that reproducibility can have different meanings in different fields of science (29–31). We explored how different approaches to reproducibility (e.g., repeatability, data sharing) were categorized by our process. We found that all papers with the MAG field of study tag “repeatability” were categorized by our method as “reproducibility” papers—in line with the National Academy of Sciences (NAS) conceptualization of reproducibility (29). Furthermore, almost all papers with the MAG field of study tags “open data” or “data sharing” were categorized by our method as “open science” papers, as intended (SI Appendix, Table S1). We should also note that the dataset for this report was compiled in 2018 (SI Appendix)—1 y before the distinction between reproducibility and replicability was formalized by the NAS report (29).

there is little difference in women's representation in leadership roles between the two literatures. Finally, we find that women's participation in high-status authorship positions is increasing over time in open science, whereas it is decreasing in reproducibility.

Taken together, we find that despite current controversies (2), the open science focus of the movement to improve science has the seed of an interconnected and prosocial culture that, if further cultivated, may continue to attract greater participation by women. We believe that the collaborative, forward-looking focus of open science has the potential to facilitate greater diversity and inclusiveness. While our focus on author gender in this article was motivated, in part, by the ability to apply validated, automated coding methods (that are highly reproducible) to determine author gender, we would nevertheless predict similar findings for scholars from other underrepresented groups. When fields are more adversarial and less prosocial, individuals from underrepresented groups (including women) may be less motivated to engage (40) due at least in part to the power dynamics described above. In contrast, fields that emphasize collaborative and prosocial norms inspire greater participation among underrepresented groups (41). It should be noted that both adversarial and collaborative cultures can engage in rigorous debate and criticism. However, collaborative cultures may afford more constructive criticism, which is a hallmark of good, forward-thinking science and what all scientists expect of peers in the field. If we wish to improve and advance the field of science, then the onus is on investigators to nurture a culture that attracts and retains a diversity of people (42–44).

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 (MAG) (45), 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 code (*Methods* and *SI Appendix*). This sample consisted of 879 open science papers and 2,047 reproducibility papers. Only 2.3% of papers shared “open science” and “reproducibility” field codes, suggesting these approaches are developing relatively independently (see *SI Appendix* for more details).

Open Science and Reproducibility Differ in Their Network Community Structures. We analyzed a total of 3,157 unique article author identification numbers (IDs) in the open science literature and 8,766 in the reproducibility literature. We built two collaboration networks using these author IDs from MAG (Fig. 1). Nodes in these networks represent scientific articles; edges represent shared authorship such that 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 interconnectedness, which suggests a more dense collaborative network within the open science literature (one-sided Fisher's exact test: $P < 0.001$).

We also performed a connected components analysis of each literature (47, 48) to measure the degree of isolation of individual subnetworks of papers within each literature (*Methods*). Results show that the reproducibility network (1,641; 0.80 components per article) contains more isolated articles (sharing fewer authors) than the open science network (661; 0.75 components per article). This components analysis indicates that the reproducibility literature's network is more fragmented. Examining the component size differences of the two networks as another indicator of connectedness, we find that the average component size (ACS) is also higher for the open science network (ACS: 1.33 vs. 1.25). Fig. 1 visualizes the two networks to facilitate interpretation of the observed network connectedness and fragmentation differences between the two literatures. 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 and smaller paper networks—and these differences between the two literatures are statistically significant ($P < 0.01$, as reported above). As a robustness check, we conducted the same analyses excluding all solo-authored papers. Results revealed that these findings are robust to this alternative analysis (see *SI Appendix* for details).

Semantic Text Analyses Suggest That the Explicit Cultures of the Open Science and Reproducibility Literatures Are Different. Using a validated text-mining dictionary (49), we measured the presence of communal and prosocial constructs (e.g., contribute, encourage, help, nurture; see *SI Appendix*, Table S2 for the list of constructs

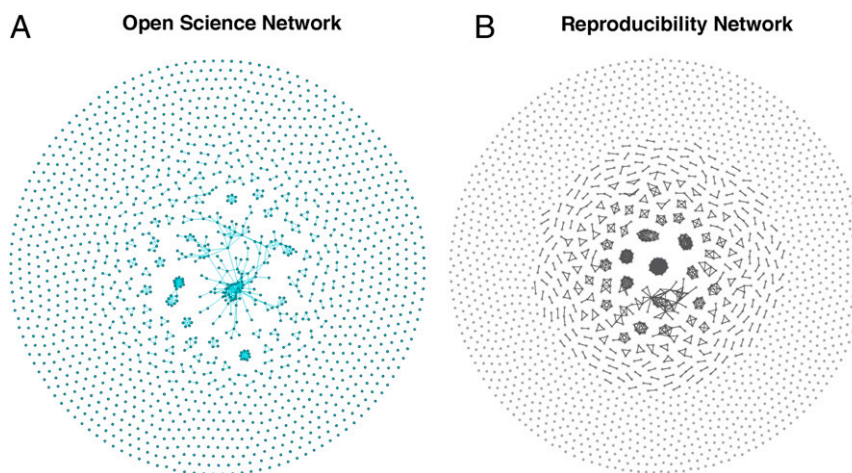


Fig. 1. Differences in author community structure: open science (A) vs. reproducibility (B). Each circle, or node, represents a scientific article. Articles share an edge (line connecting two nodes) if at least one author appears in both papers. While networks in both literatures are relatively sparse, the open science literature has formed a larger collaboration network (i.e., this community structure can be seen by the group of highly connected nodes in the center of the visualization), when compared with the reproducibility network. Data were visualized using Gephi (46).

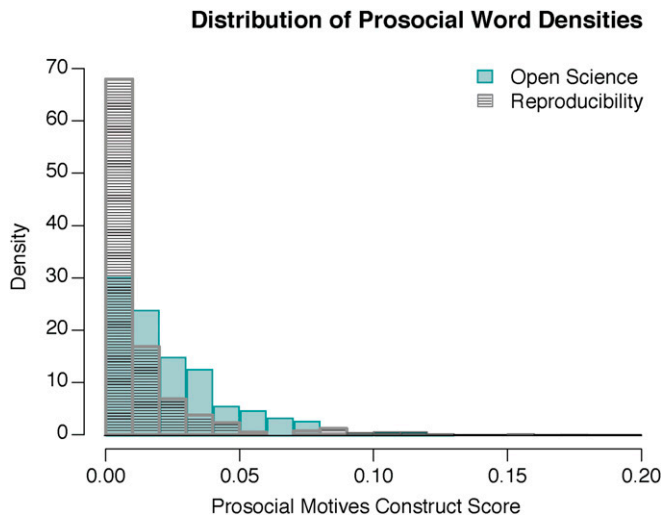


Fig. 2. Distribution of communal and prosocial word density of abstracts in the open science and reproducibility literatures. Abstracts in the open science literature include significantly more words associated with communality and prosociality than those in the reproducibility literature.

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 (two-sided test for equality of binomial proportions, $P < 0.001$). We computed the “prosocial word density” (PWD) within each dataset as the percentage of words in each abstract that reflect communal and prosocial constructs (Fig. 2 and *Methods*). The open science abstracts included more communal and prosocial words than the reproducibility abstracts (open science: mean PWD of 2.4%, median PWD of 1.8%; reproducibility: mean PWD of 0.9%, median PWD of 0.0%). A two-sided permutation test for differences in the mean and median PWD 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 PWD). Thus, we find that abstracts in the open science literature include significantly more words associated with communality and prosociality than those in the reproducibility literature.

An alternative hypothesis is that these textual differences are simply driven by disciplinary field. To examine this possibility, we stratified the model by academic field of study (i.e., computer science, engineering, medicine) and found similar effects (see *SI Appendix, Fig. S5* for details). Thus, the finding that open science incorporates more explicitly prosocial language compared with reproducibility is robust to disciplinary field.

Women’s Participation Is Differently Patterned in Open Science and Reproducibility.

Women are more likely to be represented in high-status author positions in open science (vs. reproducibility). 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. Fig. 3 displays gender representation for the open science and reproducibility literatures for single- and multiauthored papers. The single-authored subset includes 255 open science papers and 342 reproducibility papers, while the multiauthored 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 multiauthored paper.

We first analyzed single-author papers with identifiable author gender (we used an algorithm that employs census data to classify author names into the gender binary [*SI Appendix*], while acknowledging that gender is a complex and multidimensional social construct). As in scientific publishing more broadly (50–52), results revealed that, overall, women are significantly less likely than men to publish single-author papers in both literatures. An exact one-sided Binomial test indicated that the percentage of female single authors is 33.0% in the open science literature and 28.1% in reproducibility; both are lower than 50%—the proportion that would indicate gender parity ($P < 0.001$ for both tests). This suggests that women are equally engaged with each topic area in single-author roles, although underrepresented in both literatures compared with their single-author male colleagues.

For the remaining analyses, we focus on multiauthor papers. Women hold high-status authorship positions in 60.6% of the multiple-author papers in the open science literature, compared with 57.9% in the reproducibility literature. Note that with gender parity, the expected percentage of multiple-author papers with a woman in a high-status (first or last) author position would be 75% (comprised of a 25% chance of woman first and last, a

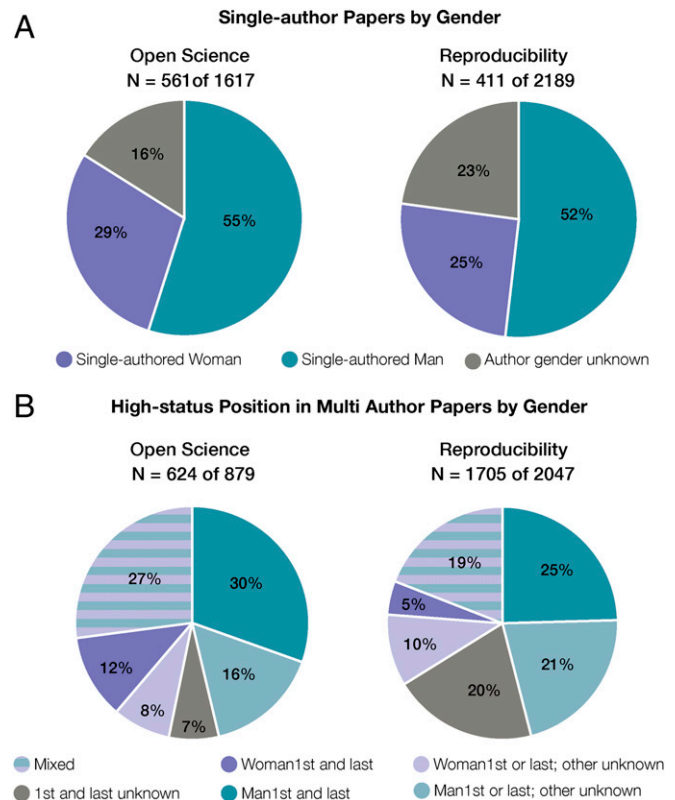


Fig. 3. Gender representation in high-status author positions (first or last) in open science and reproducibility. (A) Single-author papers by gender. Women are underrepresented in single-authored papers in both the open science and reproducibility literatures, relative to gender parity. (B) High-status positions in multi-author papers by gender. Women are underrepresented in high-status author positions in both literatures (relative to gender parity) 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) compared with the reproducibility literature (with only 34% with known female first or last author and only 5% with known female first and last author).

25% chance of woman-first and man-last, and a 25% chance of man-first and woman-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 or conference proceeding). For this analysis, we used a subset of multiauthored 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 others 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 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 multiauthor 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 ($P < 0.001$; *SI Appendix, Table S3*). Thus, whereas women are underrepresented in high-status author positions on multiauthored papers in both literatures (relative to gender parity), there is significantly greater representation of women authors in high-status author positions in the open science (vs. reproducibility) literature.

However, again, an alternative hypothesis is that these gender differences in high-status author positions are simply driven by disciplinary field. To examine this possibility, we fit the model controlling for the academic field of study and found similar effects (see *SI Appendix* for details). Thus, the gender representation difference in high-status authorship positions in open science (vs. reproducibility) is robust to disciplinary field.

Women's high-status authorship is more constrained by team size in reproducibility than in open science. Women's high-status authorship is differently patterned by team size in these literatures. Within multiauthored papers, women's likelihood of authoring in high-status positions in the open science literature is greatest in smaller teams (two- to three-author papers; Fig. 4) and remains relatively consistent as teams become larger (Fig. 5, *Left*). However, within the reproducibility literature, women are less likely to author in high-status positions in smaller teams (two- to three-author papers) and more likely to do so in larger teams (six- to seven-author papers). Regression analyses confirm this difference after controlling for other important variables, including publication year and manuscript type (Fig. 5, *Left*).

We also considered the alternative hypothesis that field differences could be driving the observed relationship between women's participation and team size. To examine this, we conducted the same regression analyses stratified by field and found that the results were largely robust across fields. That is, women are underrepresented in high-status author positions on smaller teams in the reproducibility literature (compared with the open science literature; see *SI Appendix* for a detailed description of these analyses and findings). Taken together, we find that women's participation in high-status author positions is more constrained in reproducibility than in open science and occurs more frequently in larger teams within the reproducibility literature.

Women's representation in high-status author positions is increasing in open science over time and decreasing in reproducibility. Further regression analyses reveal that in the open science literature, the representation of women in high-status authorship positions 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 ~15.6% ($P < 0.01$) year-over-year from 2010 to 2017 (*SI Appendix, Table S3*), 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 ~3.6%, although this decline 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 ($P < 0.01$). Fig. 5, *Right* illustrates the difference in trends over time between the two literatures on the probability scale.

Finally, we again explored the alternative field hypothesis: that women's participation over time was driven by field differences. Specifically, we conducted the same regression analyses stratified by field and found that the results were largely robust across fields. That is, we found growing participation of women in open science over time and decreasing participation of women in reproducibility in every field except psychology, where women's participation has grown over time in reproducibility (53) (see *SI Appendix* for detailed field analyses and findings).

Discussion

Our results reveal that the movement to improve science consists of two relatively independent groups of investigators with differing approaches: 1) open science and 2) reproducibility. These literatures have relatively few common papers and authors, indicating they are distinct, nonoverlapping communities. Each

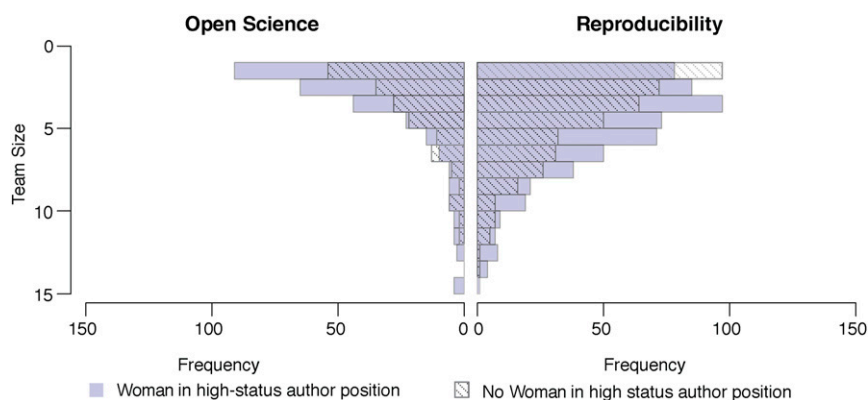


Fig. 4. Team size and women's representation in high-status positions in multiauthor papers. Women's representation in high-status authorship positions (first and last authorship) is patterned differently by team size in the open science and reproducibility literatures. Women assume high-status positions consistently across smaller and larger teams in open science, while they do so more frequently in larger teams in the reproducibility literature.

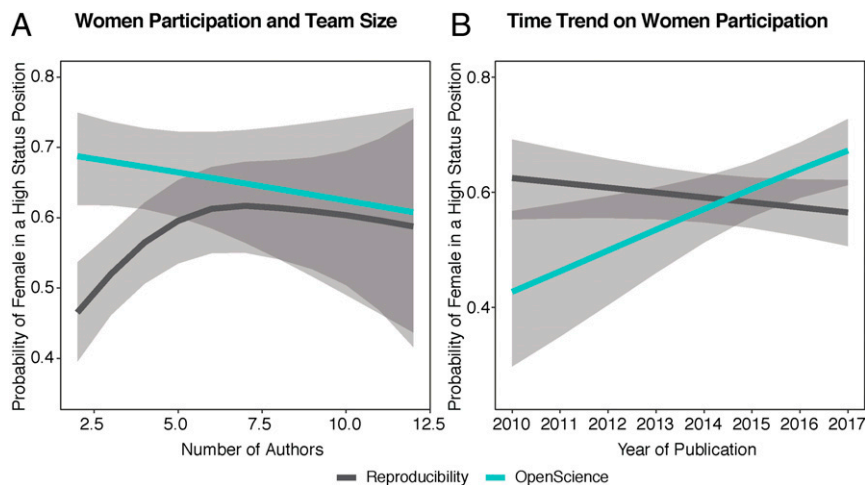


Fig. 5. Estimated regression effects of team size and year of publication on women's representation in high-status positions in multi-author papers. (A) Women participation and team size. Women have higher rates of high-status authorship in larger teams within reproducibility, while rates are comparatively and consistently high in open science across team sizes. (B) Women's 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% CIs indicated in gray. To produce the estimates, the x-axis variable and literature category are varied, while the remaining model variables are fixed (see *Methods* for details).

shows a significantly different community structure of how authors contribute to individual papers. Whereas the open science literature is significantly more interconnected with respect to coauthorship, the reproducibility literature is more fragmented. Another indicator of these different emergent cultures comes from the semantic text analysis, which suggests 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. Cohering with these structural and cultural divergences, we find different patterns of participation by women scholars. Overall, women are more likely to occupy leadership positions (i.e., high-status author positions) in the open science literature than in the reproducibility literature, and this greater participation is further patterned by team size and time. When authorship teams are relatively small (e.g., two to three authors), women's likelihood of authoring in high-status authorship positions is greater in open science compared with reproducibility. Women's participation is more constrained in reproducibility—occurring more often in larger teams—whereas it is freer in open science (occurring as frequently in smaller and larger teams). Finally, women's participation in these literatures yields different temporal patterns as well, with increasing participation in open science but decreasing in reproducibility.

Given these findings, we argue that there are strong reasons for science generally—including both subcultures of science reform—to adopt inclusive and prosocial cultures. First, a culture that portrays science as noncommunal 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 (54–56). 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 Fellowship 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. coauthored) publications (57). Today's science relies on teams coordinating their efforts to share insights and methods, build on past work, and develop new questions and approaches (58). 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) (59). 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 (60). A problem, however, is that while science is increasingly team-based, homophily processes mean that many teams are likely to be relatively homogeneous with regard to sociodemographic, behavioral, and intrapersonal characteristics (61). Attention should be paid, proactively, to the composition of teams.

Second, and consistent with the point above, 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., National Institute of General Medical Sciences 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, the National Cancer Institute's Science of Team Science Toolkit, NSF Big Data Regional Innovation Hubs 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), although there are still inequities in these processes (62).

Indeed, the complexity of the problems we are now facing in science demands the expertise of multiple disciplines working in coordinated fashion (63, 64). 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 (65, 66). In fact, data from the US 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 nature (67). Similarly, multiauthored articles are more often cited than single-authored articles (60, 68, 69), 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 (70). 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 (71, 72). In any case, if diverse team science is the future, institutions must reconsider individually constructed incentive structures as these structures may not promote rapid progress if scientists remain tied to individual incentives.

Finally, a third reason to prefer a prosocial scientific culture, consistent with our findings and that of other research, is that noncommunal practices and values may deter people who value communal, interdependent, and prosocial goals, including women (14), underrepresented minorities (41, 73), first-generation college students (73), and communally oriented men (14). 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. However, both foci have the common goal of improving our knowledge, rigor, and understanding. These contributions are likely enhanced when a diverse range of scientists are fully participating in either approach's efforts.

Lack of Diversity Can Be Problematic for Science. Lack of social diversity (e.g., gender and racial diversity) within scientific teams can be detrimental to science. There are many case studies where homogenous teams have produced serious failures of knowledge with regard to critical outcomes. For example, with no women on engineering and development teams, heart valves and seat belts are made that only fit men's bodies (significantly increasing mortality rates for women) (74), voice-recognition software only recognizes the voices of men (74), and image-recognition software tags Black people as apes (75). Including and heeding the voices and experiences of a range of people can foster outcomes that benefit a wider range of people. While teams with more gender and cultural diversity are more likely to develop new products and introduce radical innovations to market (76, 77), and while papers authored by diverse scientific teams have more citations and higher impact factors (78), the mere presence of social diversity is not always sufficient to foster equal participation of diverse social groups. 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 (79). Similarly, in gender-diverse engineering teams of students, women were underrepresented in presenting technical content, while men were overrepresented (80). Indeed, the potential of social diversity often goes untapped, leading to null or negative results on group performance (81–84).

To capitalize on the potential of social diversity, teams need to directly address the challenges that can accompany social diversity. For example, interactions and communication within diverse teams may be more difficult, especially at first (85–87). However, there is great potential of social diversity, particularly in complex tasks. Socially diverse teams encode and process information more accurately (88), especially when the sharing of disparate facts is a requirement for success (89). The mere presence of people from socially diverse backgrounds alters the cognition and behavior of majority group members to foster

improved and accurate thinking and communication (90). In the presence of social diversity, majority group members raise more facts and make fewer factual errors, and when errors are made, they are more likely to be corrected (90). 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 (91). Finally, the presence of 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 (86).

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 (92, 93). There is much to learn from this emerging movement, and several other fields (94–98) are similarly taking stock, including biostatistics (99, 100), computer science (101), and medicine (102, 103). For example, the team science approach to improving science can be observed in theoretical and experimental physics where investigative necessity has promoted large-scale consortia and successful models of scientific collaboration (104). Similarly, the collaborative discipline of structural biology established standards for sharing and deposition of code and data (see Collaborative Computational Project No. 4 and Research Collaboratory for Structural Bioinformatics), and these communal practices coincided with a broader participation of women in the field over its ten decades (105).

In sum, open science has the seed of a communal and sharing culture that, if cultivated, may continue to foster the inclusion and participation of women. We suggest that pivoting toward this cultural style could help to diversify the reproducibility movement without detracting from its core goals. 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 participate 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 (106). 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, people from lower socioeconomic backgrounds, and communally oriented men.

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, resources, and data sharing. Other specific examples include the development of Transparency and Openness Promotion Guidelines (39) and the establishment of cloud-based platforms and associated user communities for research asset sharing. See examples in PBS, data in [OpenNeuro.org](https://openneuro.org) (107), analyses in brainlife.io (108–110), and study registrations in Open Science Framework (39). Individual researchers can learn about the who, when, how, and why of their teams, including attending to the range of people represented, identifying opportunities to include diverse voices, and analyzing reasons and barriers for groups' or individuals' participation. Organizations that highlight the collaborative and communal aspects of scientific processes and success can feature connections in science, acknowledging how others help overcome stumbling blocks and rewarding teams that embody the values of open science. Each researcher can work toward broadening their collaboration and mentoring networks. 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, understanding, and action.

Methods

Data Sources. A total of 11,338 original papers were collected using the snapshot of MAG (<https://academic.microsoft.com>) on February 23, 2018. To collect the datasets, we searched MAG for all publications with specific “field of study tags” as “open science” or “reproducibility.” The field of study tags are produced by an internal Microsoft algorithm based on the contents and metadata (e.g., abstracts) of each paper (not author-generated; see ref. 111 for details). Among all of 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 included 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 published each year within each literature (*SI Appendix, Fig. S1*). As very few open science papers were published prior to 2010, and few papers in either field were published in 2018, we only use data for papers published between 2010 and 2017, which includes 2,926 papers in total, with 879 open science papers and 2,047 reproducibility papers. This is the final dataset used for all analyses, except where otherwise noted.

Data compiled for the analyses can be found at Open Science Framework (<https://osf.io/97vcx>) (112), and the code used for this work is available at GitHub (<https://github.com/everyxs/openScience>).

Based on the sample between 2010 and 2017, we constructed the paper coauthorship 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 conducted an edge density and connected components analysis as follows.

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 the edge density of 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 $7.35e^{-5}$.

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 ACS, which is defined as the network size divided by the number of connected components. The connected components analysis is conducted using the software Gephi (46).

As a robustness check, we conducted the same edge density and connected components analysis among the multiauthored papers only (excluding single-authored papers). These analyses and visualizations can be found in *SI Appendix, Fig. S2*.

Semantic Text Analysis of Abstracts. 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) and then removed those with non-English titles (79 open science and 63 reproducibility papers), as determined using the R `textcat` package (113). 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 stop words, stemming, and punctuation and converted 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 (113) (e.g., contribute, encourage, help, nurture; *SI Appendix, Table S2*). This dictionary has been shown to have acceptable agreement with human judges ($r = 0.67$) (114). The prosocial word density is calculated as the ratio of the number of prosocial words over the total number of words in each abstract. Semantic text analysis stratified by field is described in *SI Appendix, Fig. S5*.

Gender Participation Analyses. 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>) (115); 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 (noninitial) 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 US Social Security Administration baby name data from the period 1932 to 2012.

For Figs. 3 and 4, we labeled papers as having a woman 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 the 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 multiauthor papers (since in a probabilistic sense there are two “chances” to achieve high-status authorship in multiauthor papers but only one “chance” in single-author papers). In Fig. 3, we also excluded papers with more than 15 authors for the sake of visualization.

For Fig. 4, we performed logistic regression analysis to quantify how the rates of women’s high-status authorship in multiauthor papers varied by team size within each literature. We included a spline term for team size within each literature, given the evidence for a nonlinear 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 woman 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 to which each paper belongs. 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 (116). Specifically, we fit the model

$$\log\left\{\frac{\Pr(Y_i = 1)}{1 - \Pr(Y_i = 1)}\right\} = \beta_0 + \beta_1 \text{Rep}_i + \beta_2 \text{Year}_i + \beta_3 \text{Year}_i \text{Rep}_i + f_1(\text{Authors}_i) + f_2(\text{Authors}_i \text{Rep}_i) + \beta_4 \text{Conf}_i + \epsilon_i \quad \epsilon_i \sim N(0, \sigma^2)$$

where $Y_i = 1$ if paper i has a woman in a high-status author position, $\text{Rep}_i = 1$ 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 multiauthor papers), and $\text{Conf}_i = 1$ if the paper is a conference proceeding. The functions $f_1(\cdot)$ and $f_2(\cdot)$ are smooth coefficient curves that map team size to the log-odds of having a woman high-status author in each literature, given fixed values of the other coefficients.

Based on the estimated regression coefficients and SEs, we estimated the log-odds of having a woman in a lead authorship position given specific sets of predictor variables, along with normal 95% CIs. We then transformed the log-odds and CIs to odds and probabilities for better interpretability. *SI Appendix, Table S3* reports estimates and CIs on the odds scale for each parametric (i.e., nonspline) coefficient. In short, we find that the effect of belonging to the reproducibility literature is negative with an estimate of

0.393, representing ~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. Fig. 5 depicts the estimates and 95% CIs for the probability of having a female in a high-status position for different values of these variables. CIs on the probability scale are constructed by applying the inverse logit transformation [i.e., $p(x) = \exp\{x/(1-x)\}$] to the normal 95% CI on the logit scale. The estimates and CIs therefore represent predicted probabilities. In Fig. 5, *Left*, 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 woman in a high-status position. For reproducibility papers, there is a nonlinear effect of team size, with the probability of having a woman in a high-status position being markedly lower for small teams and peaking for teams with approximately seven authors before declining slightly. In Fig. 5, *Right*, we fix Conference = FALSE and Team Size = 4 (near the mean value), while allowing the 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 over time and a decline or stagnation in the reproducibility literature. Robustness checks controlling for and stratifying analyses by field are provided in *SI Appendix, Table S4* and *Figs. S3* and *S4*.

Data Availability. All data and analytic code associated with this report is publicly accessible. Data compiled for the analyses can be found at Open Science Framework (<https://osf.io/97vcx>) and code used for this work is available at GitHub (<https://github.com/everys/openScience>).

ACKNOWLEDGMENTS. M.C.M. was supported by NSF CAREER Grant DRL-1450755 and NSF Grant HRD-1661004 and by Russell Sage Foundation Grant 87-15-02. A.B.D. was supported by NSF Grant GSE-1232364. F.P. was supported by NSF Grants OAC-1916518, IIS-1912270, IIS-1636893, and BCS-1734853; NIH Grant 1R01EB029272-01; a Google Cloud Research Award; a Microsoft Investigator Fellowship; the Indiana University Areas of Emergent Research initiative “Learning: Brains, Machines, Children”; and the Indiana University Pervasive Technology Institute. We acknowledge Angela Sharpe and Cassidy Sugimoto for their 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.

1. K. Travis, The team science revolution. *Science Magazine*, 10 June 2011. <http://www.sciencemag.org/careers/2011/06/team-science-revolution>. Accessed 10 August 2020.
2. K. Finley, Diversity in open source is even worse than in tech overall. *WIRED*, 2 June 2017. <https://www.wired.com/2017/06/diversity-open-source-even-worse-tech-overall/>. Accessed 8 August 2020.
3. B. Nosek, How can we improve diversity and inclusion in the open science movement? Center for Open Science, 5 May 2017. <https://www.cos.io/blog/how-can-we-improve-diversity-and-inclusion-open-science-movement>. Accessed 8 August 2020.
4. E. A. Cech, A. Metz, J. L. Smith, K. deVries, Epistemological dominance and social inequality. *Sci. Technol. Human Values* **42**, 743–774 (2017).
5. National Academies of Sciences, Engineering, and Medicine, Policy and Global Affairs; Committee on Women in Science, Engineering, and Medicine; Committee on the Impacts of Sexual Harassment in Academia, *Sexual Harassment of Women: Climate, Culture, and Consequences in Academic Sciences, Engineering, and Medicine*, F. F. Benya, S. E. Widnall, P. A. Johnson, Eds. (National Academies Press, 2018).
6. UNESCO, *UNESCO Science Report: Towards 2030*, (UNESCO Publishing, 2015).
7. United Nations Women, Concepts and definitions. <https://www.un.org/womenwatch/osagi/conceptsanddefinitions.htm>. Accessed 8 August 2020.
8. United Nations Women, IMPACT 10x10 Initiative: Gender Parity Report 2017. <https://www.heforshe.org/sites/default/files/2018-10/HeForShe%20Gender%20Parity%20Report%202017.pdf>. Accessed 8 August 2020.
9. H. R. Markus, S. Kitayama, Cultures and selves: A cycle of mutual constitution. *Perspect. Psychol. Sci.* **5**, 420–430 (2010).
10. S. S. Luthar, Doing for the greater good: What price, in academe? *Perspect. Psychol. Sci.* **12**, 1153–1158 (2017).
11. B. Mitchneck, J. L. Smith, M. Latimer, DIVERSITY IN SCIENCE. A recipe for change: Creating a more inclusive academy. *Science* **352**, 148–149 (2016).
12. M. Syed, Why traditional metrics may not adequately represent ethnic minority psychology. *Perspect. Psychol. Sci.* **12**, 1162–1165 (2017).
13. H. L. Roediger 3rd, Varieties of fame in psychology. *Perspect. Psychol. Sci.* **11**, 882–887 (2016).
14. A. B. Diekmann, E. R. Brown, A. M. Johnston, E. K. Clark, Seeking congruity between goals and roles: A new look at why women opt out of science, technology, engineering, and mathematics careers. *Psychol. Sci.* **21**, 1051–1057 (2010).
15. G. J. Feist, Intrinsic and extrinsic science: A dialectic of scientific fame. *Perspect. Psychol. Sci.* **11**, 893–898 (2016).
16. M. A. Zárate, G. N. Hall, V. C. Plaut, Researchers of color, fame, and impact. *Perspect. Psychol. Sci.* **12**, 1176–1178 (2017).
17. A. B. Diekmann, E. K. Clark, A. M. Johnston, E. R. Brown, M. Steinberg, Malleability in communal goals and beliefs influences attraction to stem careers: Evidence for a goal congruity perspective. *J. Pers. Soc. Psychol.* **101**, 902–918 (2011).
18. A. H. Eagly, S. J. Karau, Role congruity theory of prejudice toward female leaders. *Psychol. Rev.* **109**, 573–598 (2002).
19. L. A. Rudman, C. A. Moss-Racusin, P. Glick, J. E. Phelan, “Reactions to vanguards: Advances in backlash theory” in *Advances in Experimental Social Psychology*, P. Devine, A. Plant, Eds. (Academic Press, 2012), Vol. 45, pp. 167–227.
20. M. C. Murphy, C. M. Steele, J. J. Gross, Signaling threat: How situational cues affect women in math, science, and engineering settings. *Psychol. Sci.* **18**, 879–885 (2007).
21. D. Sekaquaptewa, M. Thompson, Solo status, stereotype threat, and performance expectancies: Their effects on women’s performance. *J. Exp. Soc. Psychol.* **39**, 68–74 (2003).
22. M. C. Murphy, V. J. Taylor, “The role of situational cues in signaling and maintaining stereotype threat” in *Stereotype Threat: Theory, Process, and Application*, M. Inzlicht, T. Schmader, Eds. (Oxford University Press, 2012), pp. 17–33.
23. S. Cheryan, H. R. Markus, Masculine defaults: Identifying and mitigating hidden cultural biases. *Psychol. Rev.*, 10.1037/rev0000209 (2020).
24. M. Swers, Understanding the policy impact of electing women: Evidence from research on congress and state legislatures. *PS Polit. Sci. Polit.* **34**, 217–220 (2001).
25. M. L. Swers, *The Difference Women Make: The Policy Impact of Women in Congress*, (University of Chicago Press, 2002).
26. M. L. Swers, *Women in the Club: Gender and Policy Making in the Senate*, (University of Chicago Press, 2013).

27. M. Noland, T. Moran, B. Kotschwar, Is gender diversity profitable? Evidence from a global survey (working paper) (Petersen Institute for International Economics, 2016).
28. United Nations, The sustainable development goals report. <https://www.piie.com/publications/working-papers/gender-diversity-profitable-evidence-global-survey>. Accessed 8 August 2020.
29. H. V. Fineberg *et al.*, *Reproducibility and Replicability in Science*, (The National Academies of Sciences, Engineering, and Medicine, 2019).
30. V. Stodden, "Reproducibility in computational and data-enabled science" in *Proceedings of the 27th International Symposium on High-Performance Parallel and Distributed Computing*, (Association for Computing Machinery, 2018).
31. V. Stodden, J. Borwein, D. H. Bailey, Setting the default to reproducible in computational science research. *SIAM News* **46**, 4–6 (2013).
32. Open Science Collaboration, *PSYCHOLOGY*. Estimating the reproducibility of psychological science. *Science* **349**, aac4716 (2015).
33. B. A. Nosek, T. M. Errington, Making sense of replications. *eLife* **6**, e23383 (2017).
34. L. P. Freedman, I. M. Cockburn, T. S. Simcoe, The economics of reproducibility in preclinical research. *PLoS Biol.* **13**, e1002165 (2015).
35. R. A. Poldrack, The costs of reproducibility. *Neuron* **101**, 11–14 (2019).
36. C. G. Begley, J. P. A. Ioannidis, Reproducibility in science: Improving the standard for basic and preclinical research. *Circ. Res.* **116**, 116–126 (2015).
37. J. R. Spies, "The Open Science Framework: Improving science by making it open and accessible," PhD thesis, University of Virginia, Charlottesville, VA (2013).
38. R. A. Poldrack *et al.*, Toward open sharing of task-based fMRI data: The OpenfMRI project. *Front. Neuroinform.* **7**, 12 (2013).
39. B. A. Nosek *et al.*, SCIENTIFIC STANDARDS. Promoting an open research culture. *Science* **348**, 1422–1425 (2015).
40. M. C. Schneider, M. R. Holman, A. B. Diekmann, T. McAndrew, Power, conflict, and community: How gendered views of political power influence women's political ambition. *Polit. Psychol.* **37**, 515–531 (2016).
41. D. B. Thoman, E. R. Brown, A. Z. Mason, A. G. Harmsen, J. L. Smith, The role of altruistic values in motivating underrepresented minority students for biomedicine. *Bioscience* **65**, 183–188 (2015).
42. H. A. Valentine, F. S. Collins, National Institutes of Health addresses the science of diversity. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 12240–12242 (2015).
43. K. Intemann, Why diversity matters: Understanding and applying the diversity component of the National Science Foundation's broader impacts criterion. *Soc. Epistemology* **23**, 249–266 (2009).
44. S. E. Page, *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies*, (Princeton University Press, 2008).
45. A. Sinha *et al.*, "An overview of Microsoft Academic Service (MAS) and applications" in *Proceedings of the 24th International Conference on World Wide Web, WWW'15 Companion*, (ACM, 2015), pp. 243–246.
46. M. Bastian, S. Heymann, M. Jacomy, Gephi: An open source software for exploring and manipulating networks. *ICWSM* **8**, 361–362 (2009).
47. M. Newman, *Networks*, (Oxford University Press, 2018).
48. A.-L. Barabási, M. Pósfai, *Network Science*, (Cambridge University Press, 2016).
49. J. A. Frimer, K. Aquino, J. E. Gebauer, L. L. Zhu, H. Oakes, A decline in prosocial language helps explain public disapproval of the US Congress. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 6591–6594 (2015).
50. J. D. West, J. Jacquet, M. M. King, S. J. Correll, C. T. Bergstrom, The role of gender in scholarly authorship. *PLoS One* **8**, e66212 (2013).
51. C. W. Fox, J. P. Ritchey, C. E. T. Paine, Patterns of authorship in ecology and evolution: First, last, and corresponding authorship vary with gender and geography. *Ecol. Evol.* **8**, 11492–11507 (2018).
52. V. Larivière, C. Ni, Y. Gingras, B. Cronin, C. R. Sugimoto, Bibliometrics: Global gender disparities in science. *Nature* **504**, 211–213 (2013).
53. National Science Foundation, NSF Report: Women, underrepresented minorities gain ground in behavioral science. *APS Observer*, 25 March 2020. <https://www.psychologicalscience.org/observer/nsf-report-women-underrepresented-minorities-gain-ground-in-behavioral-science>. Accessed 8 August 2020.
54. L. Bian, S.-J. Leslie, M. C. Murphy, A. Cimpian, Messages about brilliance undermine women's interest in educational and professional opportunities. *J. Exp. Soc. Psychol.* **76**, 404–420 (2018).
55. S.-J. Leslie, A. Cimpian, M. Meyer, E. Freeland, Expectations of brilliance underlie gender distributions across academic disciplines. *Science* **347**, 262–265 (2015).
56. A. Rattan *et al.*, "Meta-lay theories of scientific potential drive underrepresented students' sense of belonging to Science, Technology, Engineering, and Mathematics (STEM)" in *J. Pers. Soc. Psychol.*, (2018), Vol. 115, pp. 54–75.
57. V. McGovern, Perspective: How to succeed in big science and still get tenure. *Science Magazine*, 31 July 2009. <https://www.sciencemag.org/careers/2009/07/perspective-how-succeed-big-science-and-still-get-tenure>. Accessed 10 August 2020.
58. D. F. Halpern, Whither psychology. *Perspect. Psychol. Sci.* **12**, 665–668 (2017).
59. S. W. Smith, IRIS—A university consortium for seismology. *Rev. Geophys.* **25**, 1203 (1987).
60. V. Larivière, Y. Gingras, C. R. Sugimoto, A. Tsou, Team size matters: Collaboration and scientific impact since 1900. *J. Assoc. Inf. Sci. Technol.* **66**, 1323–1332 (2014).
61. M. McPherson, L. Smith-Lovin, J. M. Cook, Birds of a feather: Homophily in social networks. *Annu. Rev. Sociol.* **27**, 415–444 (2001).
62. D. K. Ginther *et al.*, Race, ethnicity, and NIH research awards. *Science* **333**, 1015–1019 (2011).
63. S. Wuchty, B. F. Jones, B. Uzzi, The increasing dominance of teams in production of knowledge. *Science* **316**, 1036–1039 (2007).
64. B. F. Jones, S. Wuchty, B. Uzzi, Multi-university research teams: Shifting impact, geography, and stratification in science. *Science* **322**, 1259–1262 (2008).
65. M. L. Disis, J. T. Slattery, The road we must take: Multidisciplinary team science. *Sci. Transl. Med.* **2**, 22cm9 (2010).
66. C. Post *et al.*, Capitalizing on thought diversity for innovation. *Res. Technol. Manag.* **52**, 14–25 (2009).
67. J. Singh, L. Fleming, Lone inventors as sources of breakthroughs: Myth or reality? *Manage. Sci.* **56**, 41–56 (2010).
68. C. W. Fox, C. E. T. Paine, B. Sauterey, Citations increase with manuscript length, author number, and references cited in ecology journals. *Ecol. Evol.* **6**, 7717–7726 (2016).
69. A. Gazni, F. Didegah, Investigating different types of research collaboration and citation impact: A case study of Harvard University's publications. *Scientometrics* **87**, 251–265 (2011).
70. J. S. Katz, B. R. Martin, What is research collaboration? *Res. Policy* **26**, 1–18 (1997).
71. P. Azoulay, Small research teams "disrupt" science more radically than large ones. *Nature* **566**, 330–332 (2019).
72. L. Wu, D. Wang, J. A. Evans, Large teams develop and small teams disrupt science and technology. *Nature* **566**, 378–382 (2019).
73. J. M. Harackiewicz, E. A. Canning, Y. Tibbetts, S. J. Priniski, J. S. Hyde, Closing achievement gaps with a utility-value intervention: Disentangling race and social class. *J. Pers. Soc. Psychol.* **111**, 745–765 (2016).
74. L. Schiebinger *et al.*, Gendered Innovations in Science, Health & Medicine, Engineering, and Environment. <http://genderedinnovations.stanford.edu>. Accessed 8 August 2020.
75. W. Lee, How tech's lack of diversity leads to racist software. *SFGate*, 22 July 2015. <https://www.sfgate.com/business/article/How-tech-s-lack-of-diversity-leads-to-racist-6398224.php>. Accessed 28 February 2018.
76. C. Díaz-García, A. González-Moreno, F. José Sáez-Martínez, Gender diversity within R&D teams: Its impact on radicalness of innovation. *Innovations* **15**, 149–160 (2013).
77. M. Nathan, N. Lee, Cultural diversity, innovation, and entrepreneurship: Firm-level evidence from London. *Econ. Geogr.* **89**, 367–394 (2013).
78. R. B. Freeman, W. Huang, Collaborating with people like me: Ethnic coauthorship within the United States. *J. Labor Econ.* **33** (suppl. 1), S289–S318 (2015).
79. B. Macaluso, V. Larivière, T. Sugimoto, C. R. Sugimoto, Is science built on the shoulders of women? A study of gender differences in contributorship. *Acad. Med.* **91**, 1136–1142 (2016).
80. L. A. Meadows, D. Sekaquaptewa, "The influence of gender stereotypes on role adoption in student teams" in *Proceedings 120th ASEE Annual Conference Exposition*, (American Society for Engineering Education, Washington, DC, 2013), pp. 1–16.
81. A. D. Galinsky *et al.*, Maximizing the gains and minimizing the pains of diversity: A policy perspective. *Perspect. Psychol. Sci.* **10**, 742–748 (2015).
82. H. van Dijk, M. L. van Engen, D. van Knippenberg, 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**, 38–53 (2012).
83. A. H. Eagly, When passionate advocates meet research on diversity, does the honest broker stand a chance?: Passionate advocates and diversity research. *J. Soc. Issues* **72**, 199–222 (2016).
84. E. P. Apfelbaum, K. W. Phillips, J. A. Richeson, Rethinking the baseline in diversity research: Should we be explaining the effects of homogeneity? *Perspect. Psychol. Sci.* **9**, 235–244 (2014).
85. F. A. Rink, N. Ellemers, Temporary versus permanent group membership: How the future prospects of newcomers affect newcomer acceptance and newcomer influence. *Pers. Soc. Psychol. Bull.* **35**, 764–775 (2009).
86. N. Dasgupta, M. M. Scirle, M. Hunsinger, 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**, 4988–4993 (2015).
87. D. van Knippenberg, C. K. W. De Dreu, A. C. Homan, Work group diversity and group performance: An integrative model and research agenda. *J. Appl. Psychol.* **89**, 1008–1022 (2004).
88. S. S. Levine *et al.*, Ethnic diversity deflates price bubbles. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 18524–18529 (2014).
89. K. W. Phillips, K. A. Liljenquist, M. A. Neale, Is the pain worth the gain? The advantages and liabilities of agreeing with socially distinct newcomers. *Pers. Soc. Psychol. Bull.* **35**, 336–350 (2009).
90. S. R. Sommers, On racial diversity and group decision making: Identifying multiple effects of racial composition on jury deliberations. *J. Pers. Soc. Psychol.* **90**, 597–612 (2006).
91. A. L. Antonio *et al.*, Effects of racial diversity on complex thinking in college students. *Psychol. Sci.* **15**, 507–510 (2004).
92. A. Ledgerwood, Introduction to the special section on advancing our methods and practices. *Perspect. Psychol. Sci.* **9**, 275–277 (2014).
93. E. Yong, Psychology's credibility crisis. *Discover Magazine*, 20 January 2014. <https://www.discovermagazine.com/mind/psychologys-credibility-crisis>. Accessed 8 August 2020.
94. M. Baker, 1,500 scientists lift the lid on reproducibility. *Nature* **533**, 452–454 (2016).
95. J. S. S. Lowndes *et al.*, Our path to better science in less time using open data science tools. *Nat. Ecol. Evol.* **1**, 160 (2017).
96. D. Reilly *et al.*, Is evidence for homeopathy reproducible? *Lancet* **344**, 1601–1606 (1994).
97. M. Baker, E. Dolgin, Cancer reproducibility project releases first results. *Nature* **541**, 269–270 (2017).
98. A. Mullard, Cancer reproducibility project yields first results. *Nat. Rev. Drug Discov.* **16**, 77 (2017).
99. J. B. Buckheit, D. L. Donoho, "WaveLab and reproducible research" in *Wavelets and Statistics*, (Lecture Notes in Statistics, Springer, 1995), pp. 55–81.

100. S. Buck, Solving reproducibility. *Science* **348**, 1403 (2015).
101. M. Hutson, Missing data hinder replication of artificial intelligence studies. *Science Magazine*, 15 February 2018. <https://www.sciencemag.org/news/2018/02/missing-data-hinder-replication-artificial-intelligence-studies>. Accessed 8 August 2020.
102. A. Casadevall, F. C. Fang, Reproducible science. *Infect. Immun.* **78**, 4972–4975 (2010).
103. Open Science Collaboration, Estimating the reproducibility of psychological science. *Science* **349**, aac4716 (2015).
104. B. P. Abbott *et al.*; LIGO Scientific Collaboration and Virgo Collaboration, Observation of gravitational waves from a binary black hole merger. *Phys. Rev. Lett.* **116**, 061102 (2016).
105. S. Ward, Inspiring women in crystallography (The Cambridge Crystallographic Data Centre [CCDC], (2019).
106. C. M. Matthews, *Federal Research and Development Funding at Historically Black Colleges and Universities*, (Congressional Research Service, Library of Congress, 1993).
107. K. Gorgolewski, O. Esteban, G. Schaefer, B. Wandell, R. Poldrack, "OpenNeuro—A free online platform for sharing and analysis of neuroimaging data" in *23rd Annual Meeting of the Organization for Human Brain Mapping (OHBM) 2017*, (F1000 Research, 2017), p. 1055.
108. F. Pestilli, Human white matter and knowledge representation. *PLoS Biol.* **16**, e2005758 (2018).
109. S. Hayashi, P. Avesani, F. Pestilli, Open diffusion data derivatives. [brainlife.io. https://doi.org/10.25663/BL.P.3](https://brainlife.io/doi.org/10.25663/BL.P.3). Accessed 8 August 2020.
110. P. Avesani *et al.*, The open diffusion data derivatives, brain data upcycling via integrated publishing of derivatives and reproducible open cloud services. *Sci. Data* **6**, 69 (2019).
111. Z. Shen, H. Ma, K. Wang, "A Web-scale system for scientific knowledge exploration" in *Proceedings of ACL 2018, System Demonstrations*, (Association for Computational Linguistics, 2018), pp. 87–92.
112. F. Pestilli *et al.*, Open science, communal culture, and women's participation in the movement to improve science. *Open Science Framework*. <https://osf.io/97vcx>. Deposited 16 February 2019.
113. K. Hornik *et al.*, The textcat Package for n-Gram based text categorization in *R. J. Stat. Softw.* **52**, 1–17 (2013).
114. J. A. Frimer, N. K. Schaefer, H. Oakes, Moral actor, selfish agent. *J. Pers. Soc. Psychol.* **106**, 790–802 (2014).
115. C. Blevins, L. Mullen, Jane, John... Leslie? A historical method for algorithmic gender prediction. *DHQ* **9**, 3 (2015).
116. S. N. Wood, N. Pya, B. Säfken, Smoothing parameter and model selection for general smooth models. *J. Am. Stat. Assoc.* **111**, 1548–1563 (2016).