Earth and Space Science

RESEARCH ARTICLE
10.1029/2019EA000946

This article is a companion to Hanson et al. (2020), https://doi.org/10.1029/2019EA000930.

Association Between Author Diversity and Acceptance Rates and Citations in Peer-Reviewed Earth Science Manuscripts

J. C. Lerback[1], B. Hanson[2], and P. Wooden[2]

[1]Department of Geology and Geophysics, University of Utah, Salt Lake City, UT, USA, [2]American Geophysical Union, Washington, DC, USA

Key Points:
- Nationally and gender-diverse teams are associated with higher acceptance rates. International teams are associated with more citations
- Diversity in author teams regarding race/ethnicity in U.S.-based authorship teams is associated with lower acceptance rates and citations
- Equity, inclusion, and bias should be considered when assessing the link between diversity and manuscript outcomes

Supporting Information:
- Supporting Information S1

Correspondence to:
J. C. Lerback,
jory.lerback@utah.edu

Citation:
Lerback, J. C., Hanson, B., & Wooden, P. (2020). Association between author diversity and acceptance rates and citations in peer-reviewed earth science manuscripts. *Earth and Space Science*, 7, e2019EA000946. https://doi.org/10.1029/2019EA000946

Received 10 OCT 2019
Accepted 4 MAR 2020
Accepted article online 7 MAR 2020

Abstract

The scientific community is becoming more demographically diverse, and team science is becoming more common. Here, we compare metrics of success in academic research, acceptance rates, and citations, among/across differing team compositions regarding demographic diversity. We collected the final decisions and citations as of 2019 of 91,427 manuscripts submitted from 2012–2018 to journals published by the American Geophysical Union. We matched the authors by email on each manuscript to self-provided demographic information within the American Geophysical Union's membership database. This resulted in 20,940 manuscripts matched to nation, gender, and career stage, and 6,015 manuscripts matched to race/ethnicity for manuscripts in which the entire authorship team was affiliated with the United States.

Among similar sized authorship teams (teams of two to four), acceptance rates were 2.7, 4.5, and 0.9% higher ($p_{\text{nation}} < 0.01$, $p_{\text{gender}} < 0.05$, $p_{\text{career stage}} = 0.51$) with more than one nation, gender, and career stage, respectively, than nondiverse authorship teams. Diverse papers had 1.2 more citations for international teams than single-nation teams ($p_{\text{nation}} < 0.01$). There were 0.4 and 1.0 fewer citations for authorship teams with more than one gender or career stage than manuscripts with one gender or career stage ($p_{\text{gender}} = 0.21$, $p_{\text{career stage}} = 0.36$). Racially/ethnically diverse teams were associated with 5.5% lower acceptance rates ($p < 0.01$) and 0.8 fewer citations ($p = 0.15$) than racially/ethnically homogenous teams. Although not every difference is statistically significant, the overall results are consistent with the notion that diversity can benefit science, but equitable practices and inclusive cultures must also be fostered.

Plain Language Summary

This manuscript uses publication data from a large disciplinary scientific publisher, combined with self-reported demographic information, to understand team diversity as related to scientific outcomes. Acceptance rates and citations are used here to measure the quality of science and impact of a study on the scientific community. We find that in the case of nation, gender, and age diversity, demographically mixed teams have better outcomes. When U.S. author teams have multiracial/multiethnic teams, these scientific outcomes are lower than single-race/ethnicity teams. This is important to reinforce that diversity has the capacity to better science, but also, critically, diversity must be understood within other social contexts regarding opportunity, networks, and resource distribution.

1. Introduction

Increasingly, researchers are engaging in team-based science, including an increase in international collaborations (Jones et al., 2008; Wuchty et al., 2007). Teams have the ability to leverage the diverse experiences, tools, and perspectives each member offers (Hong & Page, 2004; Hsiehchen et al., 2015; Nielsen et al., 2017; Stahl et al., 2010). The practice of peer review is in part driven by the idea that multiple and diverse perspectives make better science (Woolley et al., 2010).

The benefits of diversity in perspective, thought, and experience are difficult to measure. In this paper, we assess multiple metrics of demographic diversity including national affiliation, gender, and career stage of authors working in teams. Additionally, we explore the race and ethnicity of U.S. authors as another metric of diversity, although this is a much smaller dataset.

We compare each type of demographic diversity to the scientific "quality" of team manuscript submissions to peer-reviewed journals to assess whether/the degree that diversity impacts science. The "quality" of

10.1029/2019EA000946

science is also difficult to measure. Acceptance rates and citations of manuscripts are useful but imperfect measures.

The effect of team diversity can be obscured in each of these “quality” metrics, but by different mechanisms, which means looking at them together provides more insight. Acceptance rate differences may reflect “quality” but also result from biases in the peer review process such as prestige or discrimination based on identifiable information of individual or a group of authors. We understand this to be an indicator of a lack of inclusive practices external to the authorship team. Differences in citations may similarly be associated with science quality but may more strongly reflect differences in social network sizes (increasing the number of colleagues who may cite) or the ability to self-promote (increasing the manuscripts' exposure).

We recognize that the diversity of a team alone does not necessarily indicate that a team will produce higher quality science. Diversity research and initiatives have been moving toward ideas of inclusion and equity over the past few decades to understand these mechanisms, and these are growing areas of research (Ahmed, 2012; Èzbilgin, 2009; Hoffman & Mitchell, 2016; Page, 2008; Pascarella et al., 2014; Roberson, 2006; Shore et al., 2011; and others). We hope that this study, by comparing multiple metrics of diversity with multiple metrics of product quality, can provide new and broader perspectives on how practices associated with diversity, equality, and inclusivity can affect science.

2. Materials and Methods

For our analysis, we used the membership and publication databases from the American Geophysical Union (AGU), the largest membership and publishing organization in the Earth and space sciences. The peer-review process at AGU is optional single-blinded, where authors often do not know the identity of their reviewers (unless they sign their reviews, but this is not tracked), but the reviewers know the name and institution of the authors. Most important, the membership data include self-identified gender (interpreted from data collected on binary sex, with an option of “Prefer not to answer”) and career stage (calculated following Ford et al., 2018 and described further in the Supplement) of more than 140,000 scientists that can be matched with the author data. The gender information is considerably more accurate that those assigned by name-nationality algorithms, and the data allow us to also consider groups that are diverse in age or career stage. It has been difficult in earlier studies to consider all of these effects together; the AGU data set makes this possible with a reasonably large sample size.

Race/ethnicity is included for scientists in the United States who opt to disclose that information. These data were collected based on the U.S. Census categories of African American, Asian American, Caucasian/White, Hispanic/Latino, Native American, Pacific Islander, Other, and Prefer Not to Answer. The “Other” category may represent multiracial or multiethnic individuals, international individuals working at U.S. institutions, or those who do not identify within the provided categories. The term “Other Minorities” in the subsequent text refers to authors who identify as African American, Hispanic/Latino, Native American, Pacific Islander, and Other.

We evaluated data for submissions and publications across all of AGU's journals from 2012–2018. This data set includes 91,427 submitted manuscripts containing 440,191 authors (non-unique). These were joined to the self-provided demographic data (Lerback & Hanson, 2017). We used the national affiliation associated with individual authors with each manuscript submission (97.4% match). In the AGU data from this period, gender is measured as a male/female binary (with an option of prefer not to declare, although we recognize that this delineation does not capture the spectrum of gender identities). We were able to match 56.6% of the authors using email addresses). We calculated career stage based on student status, graduation year of the last degree, or, lacking those, age at the time of activity (56.8% match). Race/ethnicity for U.S.-based authors is based on U.S. Census categories, 22.0% match.

In all, 18,349 authorship teams and 2,591 solo authors (22.9% of manuscripts) are fully described by nation, gender, and career stage. A total of 4,975 authorship teams and 1,040 solo authors (6.6% of manuscripts) are fully matched to race and ethnicity.

We consider different types of diverse teams representing more than one nation, continent, gender, career stage, or race/ethnicity. For this study, we use primarily the dataset where the nation, gender, and career

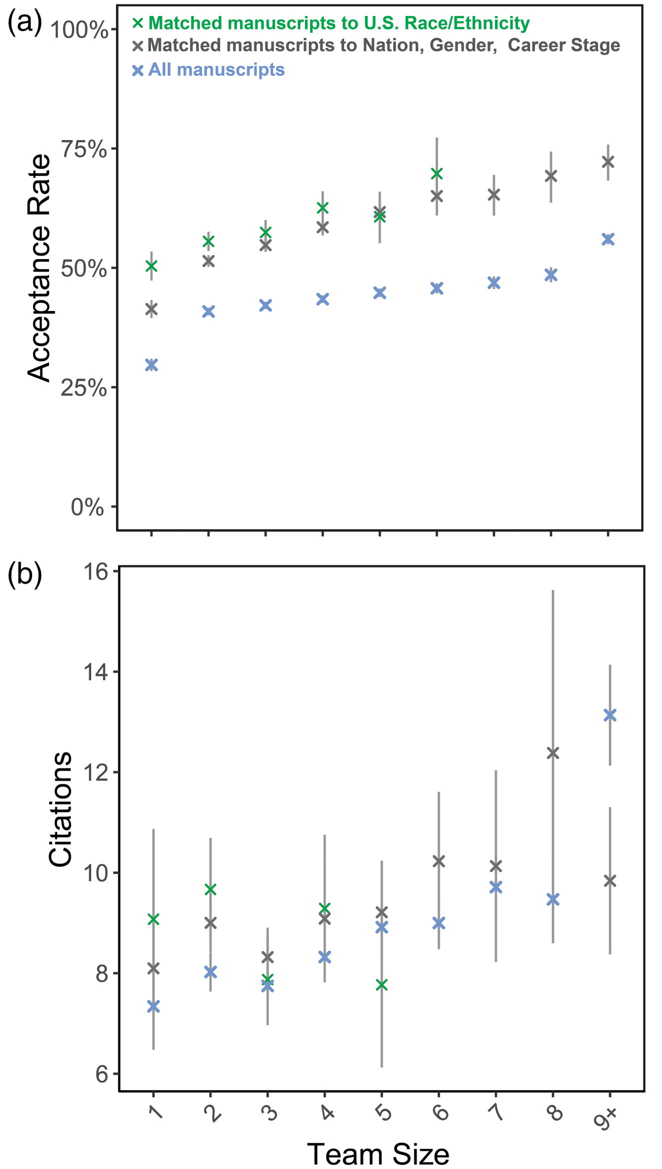

Figure 1. Authorship team size as related to scientific outcomes. (a) Acceptance rates as related to authorship team size ($n_{\text{unmatched}} = 91{,}427$, $n_{\text{matched}} = 20{,}940$). (b) Citations as related to authorship team size ($n_{\text{unmatched}} = 38{,}184$, $n_{\text{matched}} = 10{,}902$). Groups with $n < 100$ removed/error bars represent 95% confidence intervals.

stage are all matched. Where the race/ethnicity of U.S. authors are studied, we used the smaller dataset with matched race/ethnicity. Acceptance rates compare final decisions of manuscripts. Citations of 38,184 published manuscripts were collected as of early 2019 using citation counts from Clarivate Analytics and only include citations by journals indexed in the Web of Science. About 10,902 of these citation-matched manuscripts were demographically fully matched to country, gender, and career stage, and 3,262 were fully matched to race/ethnicity.

3. Data

Of all authors with nation affiliation provided, 32.9% are from the United States. Of the authors on manuscripts that are fully matched to nation, gender, and career stage information, 97.9% are affiliated with the United States. This is because international collaborations are likely to have at least one author who was not an AGU member. The next highest populations represented in this demographically matched dataset are China, the United Kingdom, France, Germany, and Canada, which make up cumulatively 1.1%. Further information on international collaborations are provided in a companion paper looking at the larger AGU Fall Meeting data set (Hanson et al., 2020). Although many of the overall international distributions are similar, the AGU Meeting attracts more authors and researchers from the United States and thus has proportionally fewer international authors and collaborations.

This dataset is generally comparable to U.S. STEM employees. Twenty-three percent of U.S. STEM employees are women (in 2018, U.S. Bureau of Labor Statistics, 2019), which is perhaps slightly underrepresented in AGU's authorship dataset where 19.3% of individuals on fully matched teams are women. U.S. STEM employees have a racial/ethnic makeup of 14% Asian American, 72% Caucasian/White, and 15% other minorities (in 2018, U.S. Bureau of Labor Statistics, 2019). Matched individuals in the AGU team-matched dataset are 13.4% Asian American, 72.6% White/Caucasian, and 14.0% other minorities (this category is described further in supporting information, similarly described by Ford et al., 2019).

Of the 18,349 authorship teams within our demographically matched dataset, 63.0% are single-nation teams, 69.1% are single-continent teams, and 89.1% of teams are single-gender. Of single-gender teams, 76.5% are all men and the other 23.5% are all women. Single-career stage teams are much less common, making up just 9.0% of teams. About 27.4% of teams are repeat teams, where the same authors submitted a different manuscript from 2012–2018 (allowing for differing author-orders). Of the racially/ethnically matched dataset ($n = 4{,}975$), 50.6% are single race teams.

4. Results

4.1. Team Size

Manuscripts with larger authors teams, regardless of diversity, tended to have been accepted at a higher rate and have more citations than manuscripts with smaller teams or single authors (Figure 1). In the matched datasets, acceptance rates increase by 17.8% in manuscripts with two to eight authors. The unmatched (full) datasets acceptance rates increase by 7.7%. Citations increase between groups of two and eight authors by 2.9 and 1.3 for matched and unmatched manuscripts, respectively. This similarity gives us some confidence that the smaller matched dataset is representative.

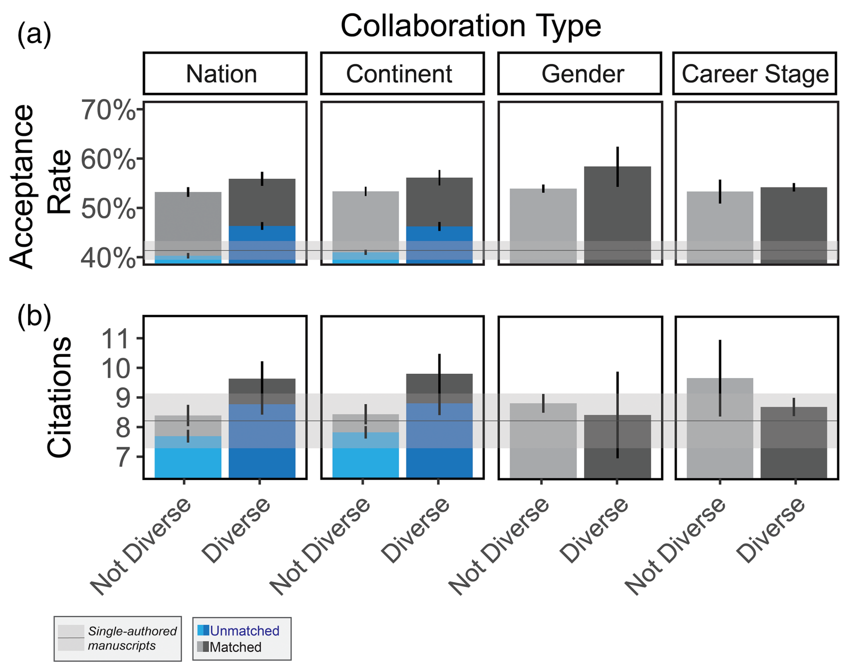

Figure 2. Types of diversity in small teams are compared to scientific outcomes. (a) Acceptance rates are compared to different team compositions with regard to international, intercontinental, gender, and career stage diversity ($n_{\text{matched small groups}} = 14{,}525$, $n_{\text{unmatched small groups}} = 45{,}179$, $n_{\text{matched single author}} = 2{,}591$). (b) Citations of manuscripts are compared to different team compositions with regard to international, intercontinental, gender, and career stage diversity ($n_{\text{matched small groups}} = 7{,}510$, $n_{\text{unmatched small groups}} = 18{,}128$, $n_{\text{matched single author}} = 1{,}007$). Error bars represent 95% confidence intervals in all figures.

This team size effect needs to be considered in assessing the importance of diversity alone because diverse teams tend to be slightly larger, due to our definition. About 79.2% ($n_{\text{manuscripts}} = 14{,}525$) of demographically matched authorship teams have group sizes of two to four authors. Thus, we focused on these smaller authorship teams of two to four members to analyze the potential effects of diversity in manuscripts to help disambiguate the effects of team size.

4.2. Nation, Gender, and Career Stage Diversity

Within these small teams fully matched for nation, gender, and career stage ($n_{\text{manuscripts}} = 14{,}525$), 67.9% of submitting teams are single-nation, and 69.1% are single-continent. About 89.1% of small teams represent single-gender collaborations. Single-gender teams were 25.0% women-only teams. Single-gender teams did not have significantly different gender-age distributions than mixed-gender teams. Teams representing a single career stage made up only 9.0% of teams. About 7,510 small teams and 1,007 single-authored manuscripts were matched to citation counts.

Diverse teams were also associated with higher acceptance rates (Figure 2a). The difference in acceptance rates for manuscripts was greater for intercontinental teams (an increase of 2.8%, $p < 0.01$) than for international teams (an increase of 2.7%, $p < 0.01$). The greatest difference (4.5%, $p < 0.05$) in acceptance rates was between single- and mixed-gender teams. There was not a significant difference between the acceptance rates between men-only (54.0%) and women-only teams (53.7%) ($p = 0.73$), nor was there a significant difference between diverse teams with a woman first-author (58.7%) and a man first-author (57.8) ($p = 0.83$). Teams made up of authors at different career stages had a 0.9% increase in acceptance rates compared to teams composed of members at the same career stage ($p = 0.51$). This difference is not statistically significant, although the direction of difference is consistent with the other tests.

Papers with international author teams generated on average 1.2 more citations ($p < 0.01$) than papers with single-nation author teams with the same number of authors (Figure 2b). This result is consistent with that of Hsiehchen et al. (2015). Manuscripts from intercontinental author teams also had more citations than single-continent teams (a 1.4 increase, significant $p < 0.01$). Papers with mixed-gender and mixed-career

10.1029/2019EA000946

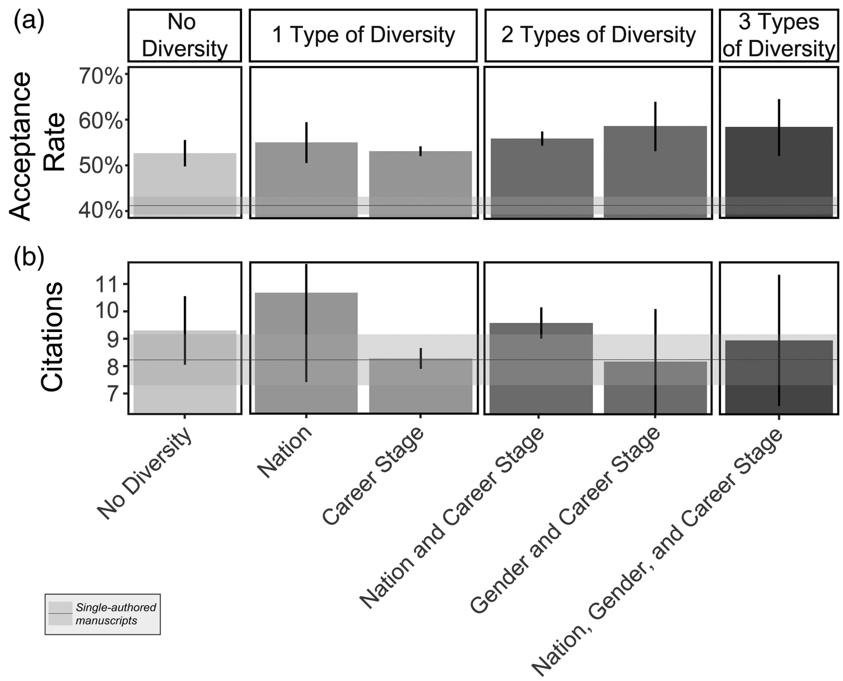

Figure 3. Intersectional diversity in small teams is compared to scientific outcomes. (a) Acceptance rates are compared to different team compositions with regard to intersectional international, gender, and career stage diversity ($n_{\text{small groups}} = 14{,}525$, $n_{\text{matched single author}} = 2{,}951$). (b) Citations of manuscripts are compared to different team compositions with regard to intersectional international, gender, and career stage diversity ($n_{\text{small groups}} = 7{,}510$, $n_{\text{matched single author}} = 1{,}007$). Collaboration types with $n < 100$ are not shown. Error bars represent 95% confidence intervals in all figures.

stage teams had slightly fewer citations on average than their nondiverse counterparts, although these differences are statistically insignificant. Mixed-gender papers had 0.4 fewer citations on average than single-gender papers ($p = 0.21$). Although nondiverse men-only teams were associated with more citations than nondiverse women-only teams (difference of 1.2 citations, $p < 0.01$), diverse teams with a first author who is a man did not have significantly more citations than diverse teams with a first author who is a woman (difference of 1.3 citations, $p = 0.43$). Mixed-career stage teams have an average of 1.0 fewer citations than single-career stage teams ($p = 0.36$).

For comparison, these results are consistent with those for the full data set where we know the country of origin of authors, but not their age or gender. Here ($n_{\text{manuscripts, small teams}} = 45{,}179$), international teams had 6.0% higher acceptance rates than single nation small teams ($p < 0.01$), although acceptance rates were ~10% lower than the matched dataset overall. International collaboration had 1.1 more citations than that of single-nation authorship teams ($p < 0.01$) ($n_{\text{cited manuscripts, small teams}} = 18{,}128$). Again, this comparison helps give confidence in the observations across the dataset with all of the authors identified also by gender and age.

Members of a team may provide multiple types of diversity (that is, one or more members can provide a combination of age, gender, and international diversity). Teams that had one or more types of diversity had generally higher acceptance rates than teams that had no diversity at all although most of these comparisons are statistically insignificant, which may in part be a reflection of the great variance in sample sizes (Figure 3a). Compared to teams with no diversity, increased citations are associated with teams that had only international collaboration ($p < 0.05$), or international and career stage collaborations ($p = 0.41$) (Figure 3b). Other combinations of diversity are either too uncommon for analysis (such as only mixed-gender collaborations) or have fewer citations overall than non-diverse papers.

4.3. Racial/Ethnic Diversity

The dataset where all team members are also identified by their race/ethnicity is small and represents 1,040 single-authored and 4,450 small-team manuscripts. Citation counts were matched to 2,943 of these

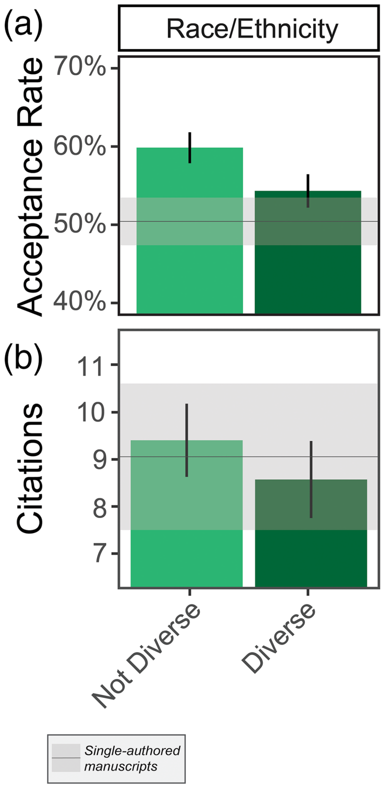

Figure 4. Small team (two to four authors) racial/ethnic diversity related to scientific outcomes. (a) Acceptance rates of manuscripts with U.S. authors are compared to different team compositions with regard to racial/ethnic diversity ($n_{\text{small U.S. groups}} = 4{,}450$, $n_{\text{matched U.S. single author}} = 1{,}040$). (b) Citations rates of manuscripts with U.S. authors are compared to different team compositions with regard to racial/ethnic diversity ($n_{\text{small U.S. groups}} = 2{,}448$, $n_{\text{matched U.S. single author}} = 495$). Error bars represent 95% confidence intervals in all figures.

(495 of these were single-author papers). Authors analyzed in this subset of manuscripts are 71.3% White/Caucasian, 12.2% Asian American, and 16.4% other American minorities. This is a small representation of the U.S. Earth and space science authorship population and applies only to racial/ethnic groups in the United States. However, quantitative studies of race/ethnicity in STEM fields are relatively uncommon, and we use this sample as an opportunity to make preliminary observations.

Of manuscripts where all authors have identified their ethnicity, 47.0% are racially/ethnically diverse. Of racially/ethnically diverse teams, 49.6% are a combination of All Other Minority and White authors, 27.3% are Asian American and White authors, 13.2% are Asian American and All Other Minority authors, and 8.9% are Asian American, All Other Minority, and Caucasian/White authors. Of the nondiverse teams, 91.3% are White/Caucasian, and 5.0% are all Asian American. There were no significant differences between the gender and career stage distributions of diverse and nondiverse teams.

Papers from racially/ethnically homogenous teams are accepted at a 5.5% higher rate than papers from racially/ethnically diverse teams ($p < 0.01$), although both types of teams fare about the same as single-authored manuscripts (Figure 4a). Manuscripts of racially/ethnically homogenous teams had an average of 0.8 more citations than diverse ones, although this difference is not statistically significant ($p = 0.15$) (Figure 4b). The opposite difference in citations has been found in other studies by Freeman and Huang (2015) and AlShebli et al. (2018). These studies, however, use an algorithm developed by Kerr (2008) and the Name Ethnicity Classifier, respectively. These classifies ethnicity by names to global ethnic categories that may be more comparable to our intercontinental or international datasets, rather than self-reported racial/ethnic identities applicable in the United States.

The experiences of people who identify with different racial groups are not the same, and these are not represented well in the racial/ethnic divisions available in the AGU dataset. Intersectionality of race/ethnicity with gender or other demographics resulted in small sample sizes and insignificant differences.

Diverse teams composed of Asian American and Caucasian/White as well as teams with Asian American, All Other Minority, and Caucasian/White authors did not have significantly different acceptance rates than single-race/ethnicity teams (differences < 3%, $p_{\text{Asian American and Caucasian/White}} = 0.749$, $p_{\text{Asian American, All Other Minority, and Caucasian/White}} = 0.443$). Teams with All Other Minority and Caucasian/White authors had 5.0% lower acceptance rate than single-race/ethnicity teams ($p < 0.01$), and teams with Asian American and All Other Minority authors had 17.3% lower acceptance rates than single-race/ethnicity teams ($p < 0.01$) (Figure 5a).

Diverse teams composed of Asian American and Caucasian/White authors had more citations than nondiverse groups, although this is not statistically significant (a difference of 1.1 citations, $p = 0.345$). Conversely, all other diverse team compositions were associated fewer citations than single-race/ethnicity teams (differences greater than one citation, $p < 0.05$) (Figure 5b). Nondiverse teams are 94.6% Caucasian/White, where other nondiverse teams each have $n < 100$ and are not reported here to preserve anonymity.

5. Discussion

To review, international teams and mixed-gender teams were associated with statistically significantly higher acceptance rates than their nondiverse counterparts. International teams were also associated with significantly more citations. Career stage diversity was not associated with a significant difference in acceptance rates or citations (although over 90% of author teams reflect career stage diversity). Mixed-race/

10.1029/2019EA000946

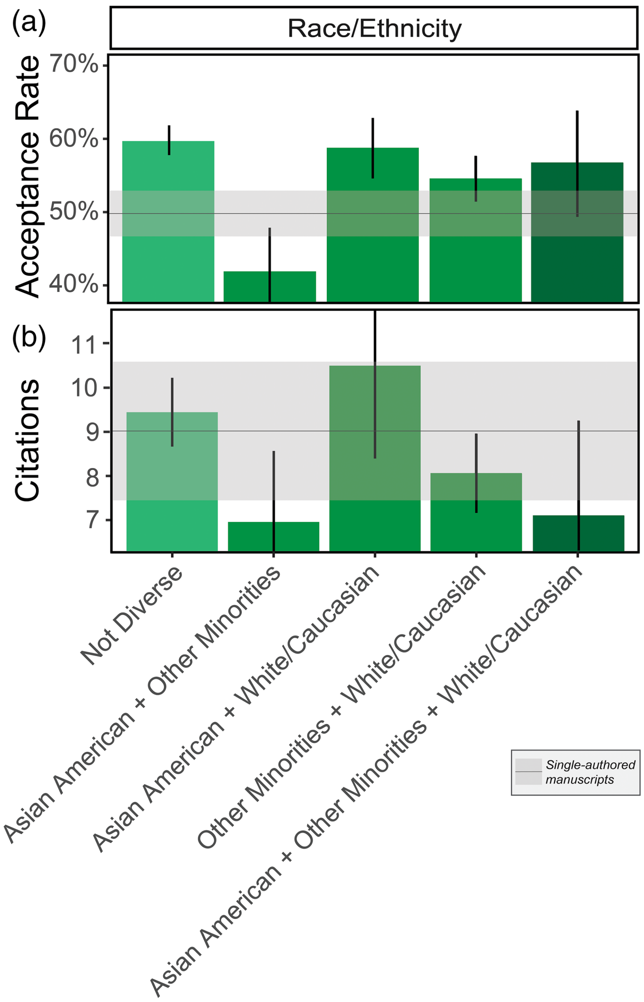

Figure 5. Racial/ethnic composition of small teams (two to four authors) related to scientific outcomes. (a) Acceptance rates of small team manuscripts with different racial/ethnic compositions ($n_{\text{small U.S. groups}} = 4{,}450$, $n_{\text{matched U.S. single author}} = 1{,}040$). (b) Citations of teams with different racial/ethnic compositions ($n_{\text{small U.S. groups}} = 2{,}448$, $n_{\text{matched U.S. single author}} = 495$). Collaboration types with $n > 100$ were removed.

ethnicity teams were associated with statistically significantly lower acceptance rates and fewer citations than racially/ethnically nondiverse teams.

The higher acceptance rates associated with national affiliation and gender may reflect diverse perspectives producing a more robust paper, making it more likely to be accepted (Woolley et al., 2010). Mixed-gender teams may have higher acceptance rates because of more effective teamwork (Leahey, 2007). Bias in the peer review process may affect these interpretations and requires further investigation (Murray et al., 2019).

More citations in international teams may reflect better science but might also be due to a network effect. If individuals have largely independent social networks (as might occur when team members are geographically separated), the manuscript might get increased exposure, enhancing the dissemination of a team paper (and possibly increase citations from) compared to a team with a smaller composite social network.

A related study by Hanson et al. (2020) shows that there are network differences across gender, national, and age-related demographic groups and that women have smaller and less international networks than men in the same age cohorts. This might reveal why mixed-gender collaborations do not necessarily result in higher citations than single-author or single-gender teams. Higher citations associated with gender-diverse teams or teams that are all men could also be due to men self-citing more often than women (King et al., 2017).

The lower acceptance rates and citations in racially/ethnically diverse teams might be attributed to a lack of inclusivity and equity and related structural/cultural support that benefits team science. Acceptance rates may be related to bias in peer-review, and citations may also reflect network effects. We recognize that this is a small dataset and hope that more data and analyses can shed more light on the specific mechanisms associated with these results.

Other research has suggested that ethnic diversity can be associated with positive effects on work produced, so our results are somewhat surprising (AlShebli et al., 2018; Cox et al., 1991; Freeman & Huang, 2015; McLeod et al., 1996). However, we know that racial/ethnic exclusion is particularly apparent in the geosciences (Bernard & Cooperdock, 2018; Dutt, 2019; Ford et al., 2019). Even in other fields of science, racial and ethnic disparities persist even when accounting for educational background, publication history, institutional, and other factors (Ginther et al., 2011, 2018). Thus, this aspect of diversity and particularly inclusion is particularly important to carefully cultivate.

Each of these results will benefit from replication and extension with more or slightly different datasets add statistical power and confidence in the results.

6. Implications

Overall, our data show that gender and international diversity can positively impact science. We believe that scientific outcomes of racial/ethnic diversity within teams reflects a culture lacking inclusivity (Bernard & Cooperdock, 2018; Dutt, 2019). Thus, these results provide an incentive for researchers to not only develop diverse author and research teams but also to consider the equitable and inclusive practices driving successful team dynamics. Diverse teams may also expand future connections, outreach, public awareness, and more inclusion for years after their formation. In addition, these results emphasize further that international collaborations and exchange of ideas benefits science and thus should be encouraged and developed rather than limited.

The differences in acceptance rates and citations that we identify highlight the importance of studying the science of research and how author teams form and interact (Cheruvelil et al., 2014). More nuanced metrics of scientific success such as the differences in citations and citing literature as developed by Wu et al. (2019) could also be used to further evaluate team dynamics. Several good examples are available around creating inclusive workplaces in order to proactively mitigate the negative effects of structural racism and other forms of exclusivity (a selected, inexhaustive list includes Holvino et al., 2004; Cho & Mor Barak, 2008; Bennett et al., 2010; Jaeger et al., 2015; Hoffman & Mitchell, 2016; Pruitt et al., 2018).

We posit that other forms of diverse life experiences associated with race/ethnicity, LGBTQ+, ableness, or other forms of identity markers should likewise be supported (where currently they are undersupported). We hope that this study will incentivize more researchers to provide race/ethnic demographic information so that AGU and similar institutions can analyze it with more statistical power and make robust, data-driven policies to better serve underrepresented populations. We encourage other societies and publishers to investigate similar questions and members to report their demographic information. We support the investigation into the magnitude of effect that other factors such as networking and homophilic biases have on the peer review and scholarly processes and how they affect acceptance rates and citations (Murray et al., 2019). Data-driven interventions can make positive changes as seen in Hanson and Lerback (2017).

Acknowledgments
We thank K. Blaufuss at AGU and the team at EJournalPress, who helped in accessing data and matching records. We thank K. Smith, B. Bowen, M. Holt, J. Holt, and A. Putman, for their comments and advice. We also thank the University of Utah's Global Change & Sustainability Center and Interdisciplinary Exchange for Utah Science research program for support in completing this research. This analysis merges a dataset on both published and unpublished AGU manuscripts with member-provided data on year of birth, gender, and personal emails. Analysis and publication of aggregated member data are consistent with AGU's privacy policy https://www.agu.org/Privacy-Policy. The matched member data are covered under this policy and cannot be released publicly. Given the information contained in and used for this analysis in the abstract data and other available information, it is not possible to fully anonymize the complete merged dataset to prevent release of any member data (even if not all of it), even if much of the data set could be assembled separately from public records. Summary tables in support of the analysis are provided in the supporting information and are hosted under the 10.5281/zenodo.3583354 (Lerback et al., 2020). Further availability of member data is possible for research under specific nondisclosure agreements by contacting AGU at AGU's discretion.

References

Ahmed, S. (2012). *On being included: Racism and diversity in institutional life*. Durham, North Carolina: Duke University Press.

AlShebli, B. K., Rahwan, T., & Woon, W. L. (2018). The preeminence of ethnic diversity in scientific collaboration. *Nature Communications*, *9*(1), 5163. https://doi.org/10.1038/s41467-018-07634-8

Bennett, L. M., Gadlin, H., & Levine-Finley, S. (2010). *Collaboration and team science: A field guide*. Bethesda, Maryland, USA: NIH Office of the Ombudsman, Center for Cooperative Resolution.

Bernard, R. E., & Cooperdock, E. H. (2018). No progress on diversity in 40 years. *Nature Geoscience*, *11*(5), 292.

Cheruvelil, K. S., Soranno, P. A., Weathers, K. C., Hanson, P. C., Goring, S. J., Filstrup, C. T., & Read, E. K. (2014). Creating and maintaining high-performing collaborative research teams: The importance of diversity and interpersonal skills. *Frontiers in Ecology and the Environment*, *12*(1), 31–38. https://doi.org/10.1890/130001

Cho, S., & Mor Barak, M. E. (2008). Understanding of diversity and inclusion in a perceived homogeneous culture: A study of organizational commitment and job performance among Korean employees. *Administration in Social Work*, *32*(4), 100–126.

Cox, T. H., Lobel, S. A., & McLeod, P. L. (1991). Effects of ethnic group cultural differences on cooperative and competitive behavior on a group task. *Academy of Management Journal*, *34*(4), 827–847.

Dutt, K. (2019). Race and racism in the geosciences. *Nature Geoscience*. https://doi.org/10.1038/s41561-019-0519-z

Èzbilgin, M. (Ed) (2009). *Equality, diversity and inclusion at work: A research companion*. Cheltenham, UK: Edward Elgar Publishing.

Ford, H. L., Brick, C., Azmitia, M., Blaufuss, K., & Dekens, P. (2019). Women from some under-represented minorities are given too few talks at world's largest Earth-science conference.

Ford, H. L., Brick, C., Blaufuss, K., & Dekens, P. S. (2018). Gender inequity in speaking opportunities at the American Geophysical Union Fall Meeting. *Nature Communications*, *9*(1), 1358. https://doi.org/10.1038/s41467-018-03809-5

Freeman, R. B., & Huang, W. (2015). Collaborating with people like me: Ethnic coauthorship within the United States. *Journal of Labor Economics*, *33*(S1), S289–S318.

Ginther, D. K., Basner, J., Jensen, U., Schnell, J., Kington, R., & Schaffer, W. T. (2018). Publications as predictors of racial and ethnic differences in NIH research awards. *PLoS ONE*, *13*(11), e0205929. https://doi.org/10.1371/journal.pone.0205929

Ginther, D. K., Schaffer, W. T., Schnell, J., Masimore, B., Liu, F., Haak, L. L., & Kington, R. (2011). Race, ethnicity, and NIH research awards. *Science*, *333*(6045), 1015–1019. https://doi.org/10.1126/science.1196783

Hanson, B., & Lerback, J. (2017). Diversifying the reviewer pool. *Eos*, *98*.

Hanson, B., Wooden, P., & Lerback, J. (2020). Age, gender, and international author networks in the Earth and space sciences: Implications for addressing implicit bias. *Earth and Space Science*, *7*, e2019EA000930. https://doi.org/10.1029/2019EA000930

Hoffman, G. D., & Mitchell, T. D. (2016). Making diversity "everyone's business": A discourse analysis of institutional responses to student activism for equity and inclusion. *Journal of Diversity in Higher Education*, *9*(3), 277.

Holvino, E., Ferdman, B. M., & Merrill-Sands, D. (2004). Creating and sustaining diversity and inclusion in organizations: Strategies and approaches. In M. S. Stockdale & F. J. Crosby (Eds.), *The psychology and management of workplace diversity* (pp. 245–276). Hoboken, NJ: Blackwell Publishing.

Hong, L., & Page, S. E. (2004). Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences*, *101*(46), 16,385–16,389.

Hsiehchen, D., Espinoza, M., & Hsieh, A. (2015). Multinational teams and diseconomies of scale in collaborative research. *Science Advances*, *1*(8), e1500211. https://doi.org/10.1126/sciadv.1500211

Jaeger, P. T., Cooke, N. A., Feltis, C., Hamiel, M., Jardine, F., & Shilton, K. (2015). The virtuous circle revisited: Injecting diversity, inclusion, rights, justice, and equity into LIS from education to advocacy. *The Library Quarterly*, *85*(2), 150–171.

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

Kerr, W. R. (2008). Ethnic scientific communities and international technology diffusion. *The Review of Economics and Statistics*, *90*(3), 518–537.

King, M. M., Bergstrom, C. T., Correll, S. J., Jacquet, J., & West, J. D. (2017). Men set their own cites high: Gender and self-citation across fields and over time. *Socius*, *3*. https://doi.org/10.1177/2378023117738903

Leahey, E. (2007). Not by productivity alone: How visibility and specialization contribute to academic earnings. *American Sociological Review*, *72*(4), 533–561.

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

Lerback, J., Hanson, B., & Wooden, P. (2020). *Data for "Association Between Author Diversity and Acceptance Rates and Citations in Peer-Reviewed Earth Science Manuscripts"*. Zenodo. https://doi.org/10.5281/zenodo.3583354

McLeod, P. L., Lobel, S. A., & Cox, T. H. Jr. (1996). Ethnic diversity and creativity in small groups. *Small Group Research*, *27*(2), 248–264.

Murray, D., Siler, K., Lariviére, V., Chan, W. M., Collings, A. M., Raymond, J., & Sugimoto, C. R. (2019). Gender and international diversity improves equity in peer review. *BioRxiv*, 400515.

Nielsen, M. W., Alegria, S., Börjeson, L., Etzkowitz, H., Falk-Krzesinski, H. J., Joshi, A., et al. (2017). Opinion: Gender diversity leads to better science. *Proceedings of the National Academy of Sciences*, *114*(8), 1740–1742.

Page, S. E. (2008). *The difference: How the power of diversity creates better groups, firms, schools, and societies—New edition*. Princeton, NJ: Princeton University Press.

Pascarella, E. T., Martin, G. L., Hanson, J. M., Trolian, T. L., Gillig, B., & Blaich, C. (2014). Effects of diversity experiences on critical thinking skills over 4 years of college. *Journal of College Student Development*, *55*(1), 86–92.

Pruitt, A. S., Brinkworth, C., Young, J., & Aponte, K. L. (2018). 5 things we learned about creating a successful workplace diversity program. *Harvard Business Review*.

Roberson, Q. M. (2006). Disentangling the meanings of diversity and inclusion in organizations. *Group & Organization Management*, *31*(2), 212–236.

Shore, L. M., Randel, A. E., Chung, B. G., Dean, M. A., Holcombe Ehrhart, K., & Singh, G. (2011). Inclusion and diversity in work groups: A review and model for future research. *Journal of Management*, *37*(4), 1262–1289.

Stahl, G. K., Mäkelä, K., Zander, L., & Maznevski, M. L. (2010). A look at the bright side of multicultural team diversity. *Scandinavian Journal of Management*, *26*(4), 439–447.

U.S. Bureau of Labor Statistics. "Labor Force Statistics from the Current Population Survey 2018–11. Employed persons by detailed occupation, sex, race, and Hispanic or Latino ethnicity" (accessed 2019): https://www.bls.gov/cps/cpsaat11.htm

Woolley, A. W., Chabris, C. F., Pentland, A., Hashmi, N., & Malone, T. W. (2010). Evidence for a collective intelligence factor in the performance of human groups. *Science*, *330*(6004), 686–688.

Wu, L., Wang, D., & Evans, J. A. (2019). Large teams develop and small teams disrupt science and technology. *Nature*, *566*(7744), 378–382. https://doi.org/10.1038/s41586-019-0941-9

Wuchty, S., Jones, B. F., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. *Science*, *316*(5827), 1036–1039. https://doi.org/10.1126/science.1136099