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] University of Utah, Department of Geology and Geophysics: 115 S 1460 E #383, Salt Lake City, UT 84112

[2] American Geophysical Union: 2000 Florida Ave. NW, Washington, DC 20009

Corresponding author: Jory C. Lerback (jory.lerback@utah.edu)

Key Points:

- Diversity in author teams regarding national affiliation, gender, and age group is associated with higher acceptance rates and citations.
- Diversity in author teams regarding race/ethnicity in U.S.-based authorship teams is associated with lower acceptance rates and citations.

Abstract

The scientific community is becoming more demographically diverse, and team science is becoming more common. Here, we compare metrics of success in STEM, such as acceptance rates and citations, between differing team compositions regarding nationality, gender, career stage, and race/ethnicity. 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 whose entire authorship team was affiliated with the U.S.

Among similar sized authorship teams (teams of 2-4), acceptance rates were 2.7, 4.5, and 0.9% higher ($p_{nation} < 0.01$, $p_{gender} < 0.05$, $p_{career\ stage} = 0.51$) with more than one nation, gender, and career stage, respectively, than non-diverse authorship teams. Diverse papers had 1.2 more citations for international teams than single-nation teams ($p_{nation} < 0.01$). 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 one career stage ($p_{gender} = 0.21$, $p_{career\ stage} = 0.36$). However, 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. These results show that diversity can have tangible benefits to science, but equitable practices and inclusive cultures must also be fostered.

Plain Language Abstract

This manuscript uses publication data from a 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 multi-racial/ethnic teams, these scientific outcomes are lower than single-race/ethnicity teams. This is important to show 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 (Wutchy 2007, Jones 2008). Teams have the ability to leverage the diverse experiences, tools, and perspectives each member offers (Hong and Page, 2004; Stahl, 2010; Hsiehchen et al., 2015; Nielsen, 2017). The practice of peer review is in part driven by the idea that multiple and diverse perspectives make better science.

However, the actual benefit of diversity of perspective, thought, and experience is 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 and assess their relationship to acceptance rate and citations of accepted papers. Additionally, we explore the self-reported race and ethnicity of U.S. authors as a metric of diversity, although this is a much smaller dataset.

Acceptance rates and citations of manuscripts are useful but imperfect measures of "quality" of science. They may indicate more creativity and the advantage of effective collaborations, but also could reflect unequal distributions of social capital or inequitable networking opportunities. For example, diverse teams with more perspectives may answer questions more robustly or engage in richer science, leading to higher acceptance rates or more citations. However, teams with more social connections and influence or resources (monetary, mentoring) may also lead to higher acceptance rates and citations than teams with less exposure or resource, leading to a bias. Through this analysis, we work to understand the distinctions between scientific quality and bias between different types of teams.

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 don't know the identity of their reviewers, but the reviewers know the name and institution of the authors. The reviewer and associate editor (used in about 35% of manuscripts sent out for peer review) may reveal their identities to the authors, but this uptake is not tracked. Most important, the membership data includes self-identified gender and birth year of more than 140,000 scientists that can be matched with the author data. Ethnicity is included for scientists in the U.S. 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 prior studies to consider all of these effects together; the AGU data set makes this possible with a reasonable large sample size.

The data on race/ethnicity 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 multi-racial or multi-ethnic 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 and 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. 4,975 authorship teams and 1,040 solo authors (6.6% of manuscripts) are fully matched to race and ethnicity.

A diverse team is defined here as 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 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. 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 U.S. 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%.

This dataset is generally comparable to U.S. STEM employees. 23% of US STEM employees are female (in 2018, U.S. Bureau of Labor Statistics, 2019, which is perhaps slightly underrepresented in AGU's authorship dataset where 19.3% of matched individuals are female. 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 dataset are 13.4% Asian American, 72.6% White/Caucasian, and 14.0% other minorities (this category is described further in the Supplement).

Of the 18,349 authorship teams within our demographically matched dataset, 63.0% are single-nation teams, and 69.1% and single-continent teams, and 89.1% of teams are single-gender. Of single-sex teams, 76.5% are all-male and the other 23.5% are all-female. Single-career stage teams are much less common, making up just 9.0% of teams. 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 and Discussion

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 homogenous teams or single authors (Fig. 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.

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. 79.2% ($n_{manuscripts}$ = 14,525) of demographically matched authorship teams have group sizes of 2–4 authors. Thus, we focused on these smaller authorship teams of 2–4 members to analyze the potential effects of diversity in manuscripts to help disambiguate the effects of team size.
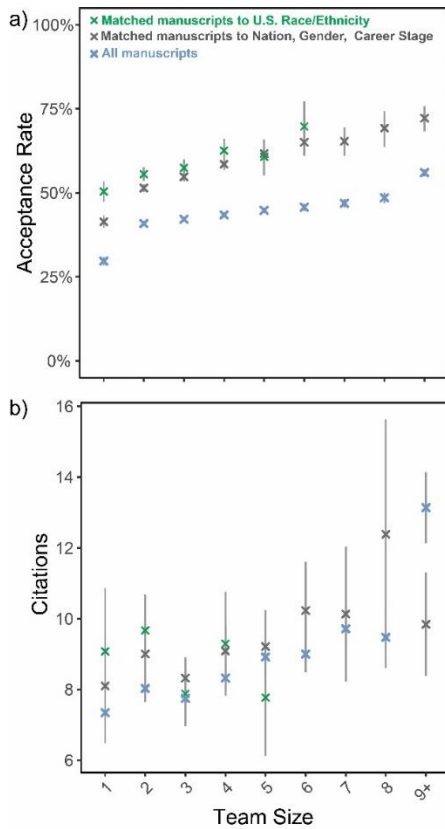

Figure 1. Authorship team size as related to scientific outcomes. a) Acceptance rates as related to authorship team size ($n_{unmatched}$ = 91,427, $n_{matched}$ = 20,940). b) Citations as related to authorship team size ($n_{unmatched}$ = 38,184, $n_{matched}$ = 10,902). Groups with n < 100 removed/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

4.2 Nation, Gender, and Career Stage Diversity

Within these small teams fully matched for nation, gender, and career stage ($n_{manuscripts}$ = 14,525), 67.9% of submitting teams are single-nation, and 69.1% are single-continent. 89.1% of

small teams represent single-gender collaborations and 9.0% of teams have authors of the same career stages. 7,510 small-teams and 1,007 single-authored manuscripts were matched to citation counts.

Diverse teams were also associated with higher acceptance rates (Fig 2.a.). 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. 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 ($p = 0.51$).

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 (Fig. 2.b.). This result is consistent with that of a previous study by 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 stage teams had slightly fewer citations on average than their non-diverse counterparts; mixed-gender papers had 0.4 fewer citations on average than single-gender papers ($p = 0.21$), and mixed-career stage teams have an average of 1.0 fewer citations ($p = 0.36$).
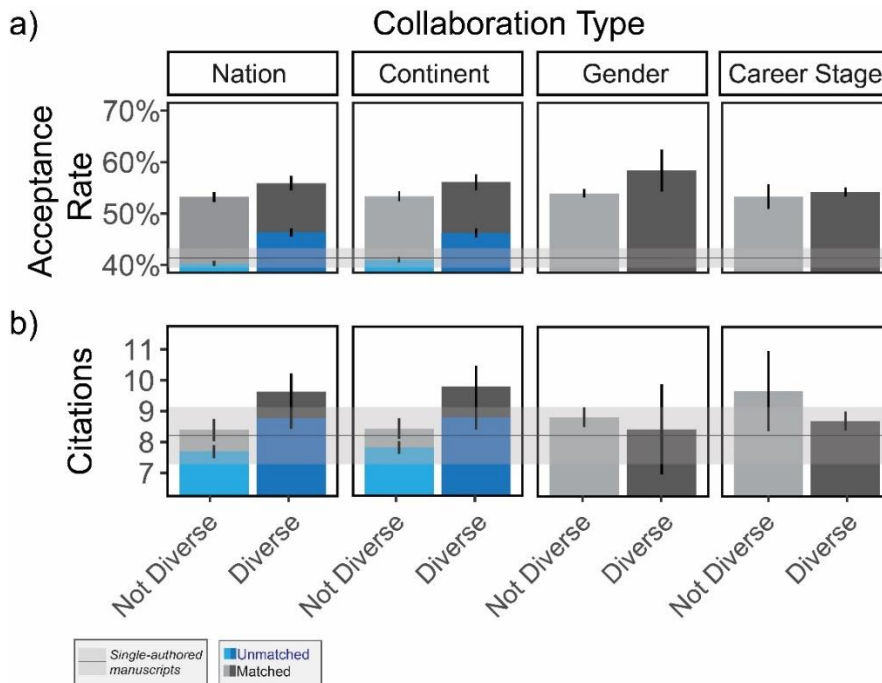

Figure 2. Types of diversity in small teams is compared to scientific outcomes. a) Acceptance rates are compared to different team compositions with regard to international, intercontinental, gender, and career stage diversity ($n_{matched\ small\ groups} = 14,525$, $n_{unmatched\ small\ groups} = 45,179$,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

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, ($n_{manuscripts,\ small\ teams} =$

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 ($p < 0.01$) ($n_{cited\ manuscripts,\ small\ teams} = 18,128$). Again, this comparison helps give confidence in the observations across the dataset with all of the authors identified also by gender and age.

If individuals in a team otherwise have largely independent social networks, this could enhance the dissemination of a team paper (and possibly increase citations from) compared to a team with a smaller composite social network. More authors are also available for self-citation in later papers. The diverse perspectives may also produce a more robust, or resilient paper, making it more likely to be accepted (Woolley et al., 2010). Mixed-gender teams may have higher acceptance rates because of more effective or diverse teamwork (Leahey, 2007). A related study by Hanson et al. (sub. 2019) shows that there are network differences by demographic groups, where women have smaller and less international networks than same-aged men. This might reveal why mixed-gender collaborations do not necessarily result in higher citations than single-author or single-gender teams.

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 higher acceptance rates that teams that had no diversity at all (Fig. 3.a.). 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 ($p = 0.41$) (Fig. 3.b.). 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.
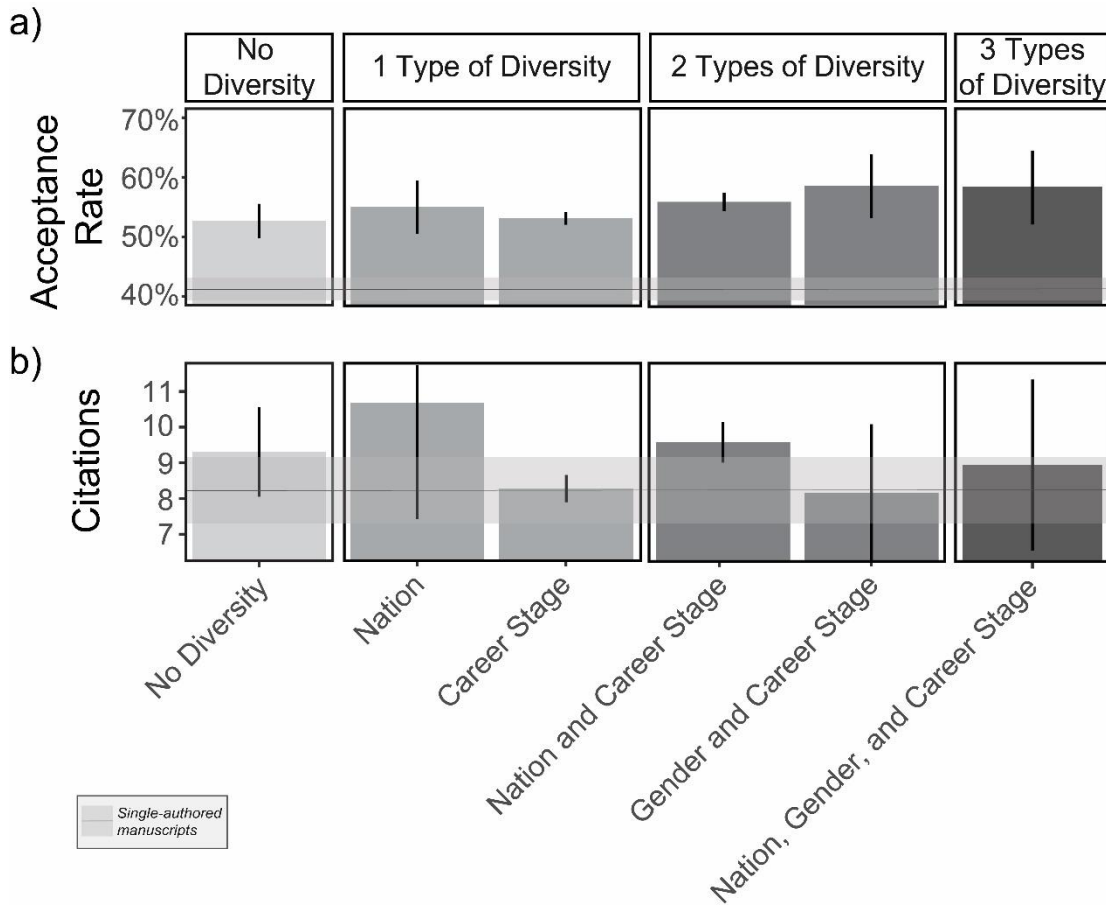
Figure 3. Intersectional diversity in small teams is compared to scientific outcomes. a) Acceptance rates are compared to different team compositions with regard to intersectional international, gender, and career stage diversity ($n_{small\ groups} = 14{,}525$, $n_{matched\ single\ author} = 2{,}951$). b) Citations of manuscripts are compared to different team compositions with regard to intersectional international, gender, and career stage diversity ($n_{small\ groups} = 7{,}510$, $n_{matched\ single\ author} = 1{,}007$). Collaboration types with $n < 100$ are not shown. Error bars represent 95% confidence intervals in all figures.

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 (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 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 U.S. 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. However, 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 (Fig. 4.a.). Manuscripts of racially/ethnically homogenous teams had an average of 0.8 more citations than

diverse ones (p = 0.15) (Fig. 4.b.). 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 U.S.
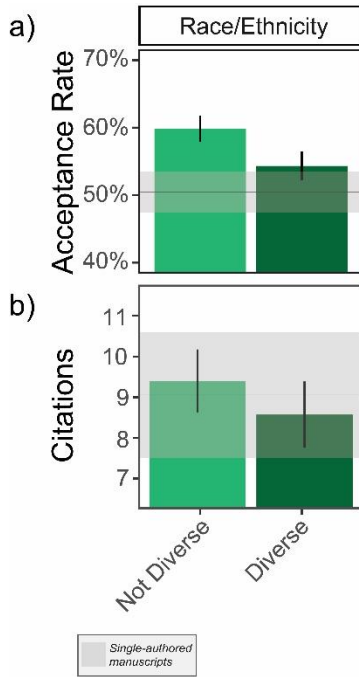

Figure 4. Small team (2-4 authors) racial/ethnic diversity related to scientific outcomes. a) Acceptance rates of manuscripts with U.S. authors are compared to different team compositions with regard to racial/ethnic diversity ($n_{small\ U.S.\ groups}$ = 4,450, $n_{matched\ U.S.\ single\ author}$ =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

The experiences of people who identify with different racial groups are not the same. Of racially/ethnically diverse teams (n = 2,091), 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. Intersectionality of race/ethnicity with gender or other demographics resulted in small sample sizes and insignificant differences.

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%, $p_{Asian\ American\ and\ Caucasian/White}$ = 0.749, $p_{Asian\ American,\ All\ Other\ Minority,\ and\ Caucasian/White}$ = 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 5.a.).

Diverse teams composed of Asian American and Caucasian/White authors had more citations than single-race 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 1 citation, $p < 0.05$) (Figure 5.b.).
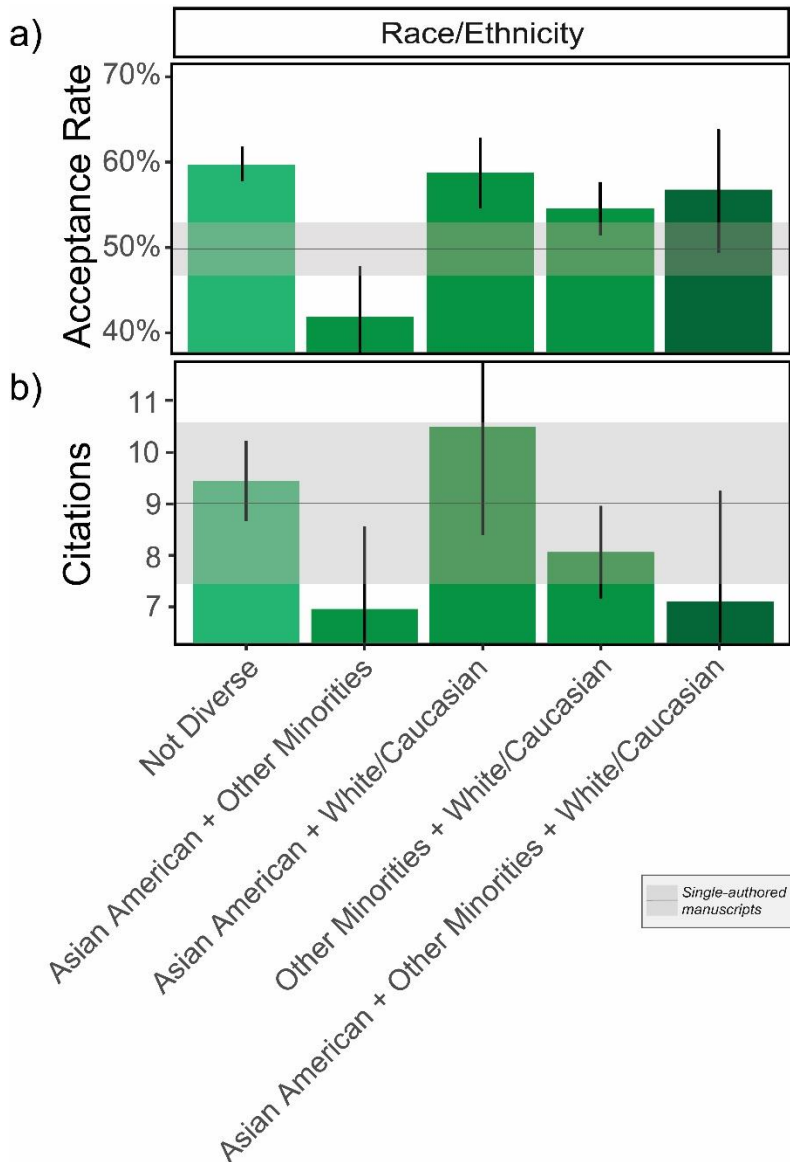

Figure 5. Racial/ethnic composition of small teams (2-4 authors) related to scientific outcomes. a) Acceptance rates of small team manuscripts with different racial/ethnic compositions ($n_{small U.S. groups}$ = 4,450, $n_{matched U.S. single author}$ = 1,040). b) Citations of teams with different racial/ethnic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

The lower acceptance rates and citations in racially/ethnically diverse teams (particularly minority-including teams) might be attributed to a variety of mechanisms. These can range from

differences in management effectiveness resulting in less creative scientific discussion and conclusions, unequal allocation of resources per capita (via geographical, institutional, or other differences), or biases in networking, team creation, or feedback received (Zenger and Lawrence, 1989; Cox et al., 1991; AlShebli et al., 2018). Racial/ethnic inclusion is particularly low (and stagnant) in the geosciences (Bernard and Cooperdock, 2018), and even in other fields, racial and ethnic disparities persist even when accounting for educational background, publication history, institutional and other factors (Ginther et al., 2011; Ginther et al., 2018).

The differences in acceptance rate and citations that we identify highlight the importance of studying the science of research and how author teams form and interact (Cheruveli et al., 2014). Lower acceptance rates and citations could be a result of bias in the peer-review process as well. 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.

5 Implications

Our data show that gender, age, and international diversity can positively impact science. 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.

We posit that other forms of diverse life experiences associated with race/ethnicity, LGBTQ+, ability, or other forms of identity markers should likewise be supported (where currently they are under-supported. 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., submitted 2018). Data-driven interventions can make positive changes as seen in Hanson and Lerback (2018).

Acknowledgments

References

1. Wuchty, S., Jones, B. F., & Uzzi, B. (2007). The increasing dominance of teams in production of knowledge. Science, 316(5827), 1036-1039.
2. Jones, B. F., Wuchty, S., & Uzzi, B. (2008). Multi-university research teams: Shifting impact, geography, and stratification in science. science, 322(5905), 1259-1262.
3. 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), 16385-16389.
4. 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.
5. Hsiehchen, D., Espinoza, M., & Hsieh, A. (2015). Multinational teams and diseconomies of scale in collaborative research. Science advances, 1(8), e1500211.
6. Nielsen, M. W., Alegria, S., Börjeson, L., Etzkowitz, H., Falk-Krzesinski, H. J., Joshi, A., ... & Schiebinger, L. (2017). Opinion: Gender diversity leads to better science. Proceedings of the National Academy of Sciences, 114(8), 1740-1742.
7. Lerback, J., & Hanson, B. (2017). Journals invite too few women to referee. Nature News, 541(7638), 455.
8. 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
9. 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.
10. Leahey, E. (2007). Not by productivity alone: How visibility and specialization contribute to academic earnings. American sociological review, 72(4), 533-561.

11. Hanson, B., Wooden, P., and Lerback, J. (Submitted as 2019EA000930, a paired paper, in 2019). Age, Gender, and International Author Networks in the Earth and Space Sciences: Implications for Addressing Implicit Bias. Earth and Space Science.

12. 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.

13. AlShebli, B. K., Rahwan, T., & Woon, W. L. (2018). The preeminence of ethnic diversity in scientific collaboration. Nature communications, 9(1), 5163.

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

15. Zenger, T. R., & Lawrence, B. S. (1989). Organizational demography: The differential effects of age and tenure distributions on technical communication. Academy of Management journal, 32(2), 353-376.

16. 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.

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

18. 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.

19. 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.

20. 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. Doi: 10.1890/130001

21. Wu, L., Wang, D., & Evans, J. A. (2019). Large teams develop and small teams disrupt science and technology. Nature, 566(7744), 378.

22. 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.

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