UC Riverside
UC Riverside Previously Published Works

Title

Does the Match between Gender and Race of Graduate Teaching Assistants and Undergraduates Improve Student Performance in Introductory Biology?

Permalink

https://escholarship.org/uc/item/3283p9rm

Journal

Authors

Lee, Star

Ing, Marsha

Publication Date

DOI

Peer reviewed

Does the Match between Gender and Race of Graduate Teaching Assistants and Undergraduates Improve Student Performance in Introductory Biology?

Star W. Lee†* and Marsha Ing‡

†Department of Developmental and Cell Biology, University of California, Irvine, Irvine, CA 92697 and ‡Graduate School of Education, University of California, Riverside, Riverside, CA 92521

ABSTRACT

Research suggests there are potential benefits to students when taught by instructors who share the same gender and/or race/ethnicity. While underrepresented students have shown increased persistence and academic performance when they were taught by gender- and/or race/ethnicity-congruent faculty, there is little research that has explored the influence of matching for graduate student teaching assistants (GTAs). Given that science, technology, engineering, and mathematics (STEM) GTAs spend a significant amount of time with undergraduates, measurable impacts on student outcomes have the potential to contribute to the success of undergraduates who have been underrepresented in STEM fields. This study evaluated the effects on academic performance of GTA ($n = 50$) matching for first-year students ($n = 976$) in an introductory biology lab course at a Hispanic-serving institution. There was no significant difference in academic performance for students who matched with the gender, race/ethnicity, income, and first-generation status of their GTAs. Results were consistent across multiple cohorts of students, after including statistical controls for prior academic performance and other demographic characteristics and accounting for the nested structure of the data. These results suggest there is a need of supporting GTAs to develop more effective teaching practices and to consider effects of GTA matching on other outcomes.

INTRODUCTION

Over 40 years (1971–2012), there was an increase of about 10% in the number of incoming full-time students interested in pursuing a major in science, technology, engineering, or mathematics (STEM; Eagan *et al.*, 2014). However, there are gender and race/ethnicity differences in the students who complete their degrees. Females and students who identify as American Indian, Alaska Native, Black/African American, Hispanic/Latinx, Native Hawaiian, and other Pacific Islander (referred to as underrepresented minorities [URM]) are consistently underrepresented in STEM fields (National Academy of Sciences, National Academy of Engineering, and Institute of Medicine, 2011; National Science Foundation, National Center for Science and Engineering Statistics, 2019).

Although female and URM students enroll at 4-year institutions with the intent to major in STEM fields, they are less likely to graduate with STEM degrees (Eagan *et al.*, 2014; Estrada *et al.*, 2016). One contributing factor to persistence and academic performance is that URM students perceive the STEM environment as less accepting and have a reduced sense of belonging in STEM compared with male and White students (Seymour and Hewitt, 1997; Walton and Cohen, 2011; Good *et al.*, 2012; Johnson, 2012). Female students of color in STEM experience bias and are not recognized as scientists by others (Carlone and Johnson, 2007; Johnson, 2012; Ong *et al.*, 2011). Among other factors, being validated, experiencing a positive racial climate, and

positive interactions with faculty are correlated with increased persistence and sense of belonging among female and URM students (Hurtado and Carter, 1997; Maestas *et al.*, 2007; Johnson, 2012; Clark *et al.*, 2016; Alcantar and Hernandez, 2020). Quality interactions between students and faculty are a predictor of student learning for all race/ethnic groups and are more impactful for students of color compared with White students (Lundberg and Schreiner, 2004).

The frequency of student–faculty interactions varies depending on the race/ethnicity of the students and the type of institution. Colleges and universities may identify as serving specific communities of students, depending on their histories, goals, or student demographics. Historically black colleges and universities (HBCU) are institutions that were established before 1964 with the purpose of educating Black students (National Academies of Sciences, Engineering, and Medicine, 2019). Black students at HBCUs are more likely to interact with faculty compared with their Black peers attending predominantly white institutions (PWIs; Hurtado *et al.*, 2011; Nelson Laird *et al.*, 2007). The interactions may be due to students being able to find same-race role models more easily at an HBCU than a PWI (Hurtado *et al.*, 2011), with more than 50% of faculty being Black at a HBCU compared with less than 5% at a PWI (National Academies of Sciences, Engineering, and Medicine, 2019). However, the increase in student–faculty interactions does not apply to Hispanic/Latinx students at Hispanic-serving institutions (HSIs; Nelson Laird *et al.*, 2007; Hurtado *et al.*, 2011). HSIs are defined as institutions with at least 25% of full-time enrolled students being Hispanic/Latinx (National Academies of Sciences, Engineering, and Medicine, 2019). Hispanic/Latinx students at HSIs interact with faculty comparably to Hispanic/Latinx students at PWIs (Nelson Laird *et al.*, 2007; Hurtado *et al.*, 2011). While there are more Hispanic/Latinx faculty at HSIs than PWIs, only 23–44% of all faculty identify as Hispanic/Latinx at HSIs (National Academies of Sciences, Engineering, and Medicine, 2019), and less than 15% of STEM faculty at HSIs are Hispanic/Latinx. The underrepresentation of minority faculty members is of concern, as it may contribute to the reduced frequency and quality of interactions between faculty and URM students.

Increasing the diversity of instructors may have a positive impact on academic performance, confidence, choice of major, and persistence, specifically when race/ethnicity or gender of the instructors and students are congruent (Robst *et al.*, 1998; Bettinger and Long, 2005; Hagedorn *et al.*, 2007; Hoffmann and Oreopoulous, 2009; Carrell *et al.*, 2010; Cotner *et al.*, 2011; Eddy *et al.*, 2014; Fairlie *et al.*, 2014). With more Latinx faculty on campus, Latinx students perform better academically (Hagedorn *et al.*, 2007). Having a URM instructor increases the likelihood of URM students passing classes and reduces the minority opportunity gap in community college courses (Fairlie *et al.*, 2014). Similarly, matching the gender of instructors and students results in better grades in a course, and students are less likely to drop the course (Hoffmann and Oreopoulous, 2009). The gender gap is reduced when female students are taught by female professors in math and science courses (Carrell *et al.*, 2010; Eddy *et al.*, 2014; Bailey *et al.*, 2020). With female instructors, female students are more confident and more likely to major in a given subject and persist into their second year (Robst *et al.*, 1998; Bettinger and Long, 2005; Cotner *et al.*, 2011). The potential benefits to students when instructors match their gender or race/ethnicity may be a result of instructors serving as role models; cultural synchronicity, wherein sharing cultural backgrounds may increase culturally relevant pedagogy; and higher instructor expectations for their students (Villegas and Irvine, 2010).

The benefits of instructor matching may depend on a number of student factors, including gender, race/ethnicity, discipline, and year of study. The impact of female faculty on female students' persistence in science may be greater when there is a lower percentage of female students enrolled in a course (Robst *et al.*, 1998). Depending on the discipline, female instructors influence the likelihood of female students enrolling in a subsequent course; there is a positive effect in geology, mathematics, and statistics and a negative effect in biology and physics (Bettinger and Long, 2005). First-year students may perceive the absence of role models as more of an issue and prefer mentors of the same race/ethnicity or gender compared with older students (Hartman and Hartman, 2008; Blake-Beard *et al.*, 2011). It is more important for female and URM students to have mentors of the same gender or race/ethnicity, who provide increased support, compared with male and White students (Blake-Beard *et al.*, 2011).

In addition to faculty, the graduate student population represents a potential pool of role models or mentors for undergraduate students. Graduate teaching assistants (GTAs) have more opportunities to interact individually with undergraduate students in smaller classroom settings compared with faculty in larger lecture halls. GTAs are responsible for the majority of the laboratory or discussion sections for biology courses at many research universities and are typically the sole instructors for their lab or discussion sections (Rushin *et al.* 1997; Sundberg *et al.*, 2005). Latinx graduate students at an HSI recognize and feel a sense of responsibility to mentor Latinx undergraduate students (Marin and Pereschica, 2017). An increased percentage of female and URM STEM graduate students is associated with increased persistence of female and URM undergraduate students (Griffith, 2010).

The impact of GTA-matching gender or race/ethnicity on undergraduate student academic performance in STEM courses is largely unknown. In a population of predominantly White GTAs and White undergraduate students, gender or race/ethnicity of the GTA did not predict student learning in a laboratory course (Wheeler *et al.*, 2017). However, it is unclear whether female and/or URM students would benefit academically by being taught by GTAs who matched their gender or race/ethnicity at an HSI.

This study addresses the following research questions:

1. How does the GTA population match the undergraduate student population in terms of demographics (gender, race/ethnicity) at an HSI?
2. Do students perform better in a biology laboratory course when matched with one demographic characteristic of the GTA? This question investigates the benefits of matching for students of any gender and race/ethnicity.
3. More specifically, do female and Latinx students perform better in a biology laboratory course when taught by gender- or race/ethnicity-congruent GTAs? This question investigates the benefits of matching in specific subpopulations, including only females, Latinx, or Latina students.

TABLE 1. First-year students' demographics in percentages[a]

| | Fall 2014 | | Fall 2015 | | Fall 2016 | | Fall 2017 | |
	All (n = 4279)	STEM majors (n = 1574)	All (n = 4029)	STEM majors (n = 1506)	All (n = 5358)	STEM majors (n = 1995)	All (n = 4599)	STEM majors (n = 1740)
Gender								
Female	55	56	54	56	58	58	56	56
Male	45	44	46	44	42	42	44	44
Race/ethnicity								
Asian	40	45	39	43	30	33	36	39
Black	4	3	3	3	3	3	3	2
Latinx	36	32	37	34	48	44	41	38
White	12	13	11	11	10	11	11	11
Other	8	7	10	9	9	9	9	10
First generation								
No	45	47	47	57	38	41	47	50
Yes	55	53	53	49	62	59	53	50
Low income								
No	55	58	55	58	49	51	56	59
Yes	46	42	45	42	57	49	44	41

[a]Source: Institutional Research Office.

MATERIALS AND METHODS

Institutional Background and Administrative Data

This study used undergraduate and graduate student administrative data from a large public university with a Carnegie basic classification of Doctoral University: Highest Research Activity. Participants were enrolled at an HSI (38% of the undergraduate student population identifying as Hispanic or Latinx). All were first-year students entering in Fall 2014, Fall 2015, Fall 2016, or Fall 2017 (Table 1). On average, ~41% of the undergraduates were Hispanic or Latinx and 37% of the first-year STEM majors were Hispanic or Latinx. The study was approved by the institutional review board (protocol number HS-19-090).

The data used in the study consisted of demographic information of both undergraduate students and their respective introductory biology laboratory GTAs. Demographic information included gender and race/ethnicity. Gender and race/ethnicity were self-reported. Gender was categorized as: female, male, nonbinary. Race/ethnicity data were categorized in the following terms: American Indian/Alaska Native, Asian, Black/African American, Hispanic/Latinx, Native Hawaiian or other Pacific Islander, White, non-resident alien, unknown, multiracial. Due to the small sample size for students in the race/ethnicity categories reported in Table 1, American Indian/Alaska Native, Native Hawaiian or other Pacific Islander, non-resident alien, unknown, and multiracial were combined to create an "Other" category. Multiracial referred to any combination of ethnicities. Subdividing among multiracial resulted in a low number of matches between undergraduates and GTAs. Similarly, non-resident aliens and unknowns provided no additional information to be able to match ethnicities between undergraduate students and GTAs. Thus, undergraduate students and GTAs categorized as multiracial (n = 97), non-resident alien (n = 17), or unknown (n = 27) were excluded from this study. In addition, students who identified as American Indian or Alaska Native (n = 1) and Native Hawaiian or Other Pacific Islander (n = 3) were excluded from the study based on the small samples in each population and no chance of matching within the

GTA population. Other demographic information for undergraduates included whether or not the student was a first-generation college student and whether or not the student was considered to fall within the low-income category. These variables were used as covariates in the analyses. In addition to demographic information, academic information from undergraduates such as prior academic performance as measured by the Scholastic Aptitude Test (SAT) or ACT were included. If students did not take the SAT but took the ACT, then their ACT scores were converted to a scale similar to the scale used for the SAT.

We analyzed students' course grades from an introductory biology laboratory course during Fall 2014 to Spring 2018 quarters. Grades from the course were converted and analyzed on a 4.0 scale ("A/A+" = 4.0, "A–" = 3.7, "B+" = 3.3, etc.) to allow for quantitative data analysis. Figures report grades as categorical: "A", "B", "C", "D", and "F". For example, in Figure 1, the grade "A" category includes "A+", "A", and "A–". All academic performance data were maintained by the administrative database (not self-reported).

More specifically, this study included only undergraduate students who were first-year freshmen, aged 18 and over, and taking the course for the first time. There were 14 undergraduates (<1%) who enrolled in the course more than one time who were dropped from the analyses. The reason for excluding these students is that we could not isolate the impact of GTA match, because each student had multiple GTAs for the same course. In addition, the sample included only students who were majoring in the life sciences (biochemistry; biology; entomology; neuroscience; plant biology; cellular, molecular, and developmental biology; microbiology; and undeclared in the life sciences; n = 976).

All undergraduates had experience with only one GTA in their introductory biology courses. The GTAs (n = 50) were included in the study if they taught the introductory biology laboratory course at least once. All GTAs were full-time graduate students pursuing doctoral degrees in the life sciences. All GTAs were required to complete a 1-day, campus-wide TA orientation that focused on the following topics: undergraduate resources;
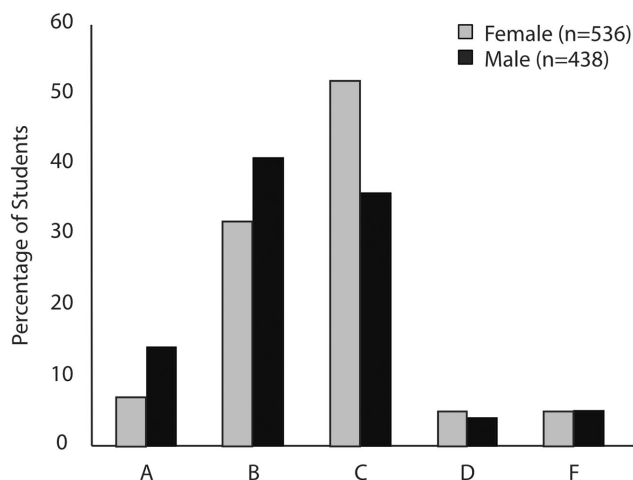
FIGURE 1. Grades in the introductory biology lab course by gender. The percentage of male students who received an "A" or "B" was greater than the percentage of female students. Fewer percentage of male students received a "C" compared with female students. Male students, M = 2.54, SE = 0.04, performed statistically higher than female students, M = 2.32, SE = 0.04; $t(972) = 3.87$, $p < 0.001$. Note: Each grade includes "+" and "−". For example, "A" includes a grade of "A", "A+", and "A−". Gender information was missing for two students. Course grades for these students were not included.

responsibilities, resources, and rights of GTAs; rapport and diversity; and organization, knowledge, and instruction.

Compared with undergraduate students, GTAs were more likely to be male, $t(88) = 8.48$, $p < 0.001$, and White, $t(88) = 19.54$, $p < 0.001$, and less likely to be low income, $t(88) = −27.97$, $p < 0.001$, or first generation, $t(88) = 43.24$, $p < 0.001$ (Table 2).

Introductory Biology Laboratory Course
All life science majors were required to take a series of three introductory biology courses. In the first quarter, the students were required to enroll concurrently in two separate courses: a lecture course and a laboratory course. Both of these courses were focused on cellular and molecular biology. In each course,

TABLE 2. Percentage of undergraduate students (n = 976) and GTAs (n = 50)

	Undergraduate	GTA
Gender		
Female	55	36
Male	45	64*
Race/ethnicity		
Asian	52	28
Black	3	—[a]
Latinx	33	12
White	12	60*
Low income	21	14
First generation	26	6

[a]Excluded due to low sample size (n = 1).
*$p < 0.001$.

students received a grade that was independent of the other course. The study focused on the laboratory course, because GTAs were the only instructors interacting with students each week.

GTAs met weekly with an academic coordinator to discuss the upcoming week's lesson plan. GTAs were the sole instructors in the laboratory for 24 students per section, and each GTA taught two sections per quarter. The laboratory course met for 3 hours once a week for 10 weeks. Students enrolled in a lab section without any information about the GTAs assigned to their sections. GTAs were assigned to their two lab sections the week before classes started.

Each week, the cellular and molecular biology laboratory course focused on one of the following topics: scientific method, microscopy, diffusion and osmosis, spectrophotometry, enzymatic reactions, fermentation and respiration, transformation, and polymerase chain reaction. Students worked in groups of two to four students. In each lab, students completed a worksheet, which consisted of knowledge, comprehension, and application-level questions. Summative assessments in the course included three quizzes and a final exam, which consisted of knowledge, comprehension, and application-level questions.

Student Performance by Gender. Figure 1 illustrates percentages of undergraduate student grades in the introductory biology lab course by gender. Overall, the percentage of male students who received an "A" or "B" (14 and 41%, respectively) was greater than that of female students (7 and 32%, respectively); whereas 36% of male students received a "C" compared with 52% of female students. Males (M = 2.54, SD = 0.94) performed statistically higher than females, M = 2.32, $SD = 0.84$; $t(972) = 3.87$, $p < 0.001$. This pattern was consistent each year.

Student Performance by Race/ethnicity. Figure 2 illustrates percentages of undergraduate student grades in the introductory
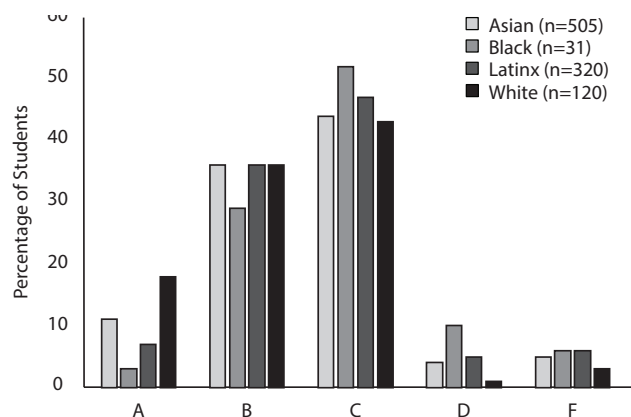
FIGURE 2. Grades in the introductory biology lab course by race/ethnicity. There were overall differences in course grades by race/ethnicity, $F(3, 972) = 5.44$, $p < 0.01$. Based on Tukey's post hoc comparisons, white students (M = 2.68, SE = 0.08) received significantly higher grades than Black (M = 2.16, SE = 0.15, Tukey HSD = 5.43, $p < 0.05$) and Latinx (M = 2.33, SE = 0.05, Tukey HSD = 3.69, $p < 0.05$) students. There were no differences in grades for White students compared with Asian (M = 2.44, SE = 0.04) students. Note: Each grade includes "+" and "−". For example, "A" includes a grade of "A", "A+", and "A−".

biology lab course by race/ethnicity. The percentage of White students who received an A is greatest (18%) followed by Asian (11%), Latinx (7%), and Black (3%) students. There were overall differences in course grades by race/ethnicity, $F(3, 972) = 5.44$, $p < 0.01$. Post hoc comparisons indicate that White students performed statistically higher than Black (Tukey HSD test = 5.43) and Latinx students (Tukey HSD test = 3.69).

Ordinary least-squares regression models were used to determine whether students' grades were related to GTA and undergraduate matching in terms of gender, race/ethnicity, or low-income or first-generation status. The models control for the demographic and prior academic performance covariates described earlier. Our variables of interest are dichotomous variables that indicated whether students matched with their introductory biology GTAs in terms of the different demographic characteristics. If there was a match, the students received a score of 1. If there was no match, the student received a score of 0. For example, a male undergraduate who had a male GTA would receive a 1 for the gender match variable. The same undergraduate might be Latinx and the GTA would be White and the undergraduate would receive a score of 0 for the match race/ethnicity match variable. These match variables were created for each characteristic (gender, race/ethnicity, low-income status, first-generation status) and entered into the regression models separately. There were four regression models: the gender match variable and covariates predicting course grades; the race/ethnicity match variable and covariates predicting course grades; the low-income match variable and covariates predicting course grades; and a first-generation status match variable and covariates predicting course grades. We checked for assumptions of linearity, homoscedasticity, and normality. We also checked for multicollinearity of the predictor variables by examining the variance inflation factors (all values < 5 across all four models). These data violate the independence assumption, because the observations are not completely independent. Students within the same labs with the same GTA have more shared experiences than students in a different lab with a different GTA.

The model-building process included trying to understand the data in terms of differences by year, by quarter, by lab section, by gender, by ethnicity, and so one, all of which were not reported. For example, we ran random effects models like multilevel regression models, which was another approach to account for the nonindependence of our level-1 units (students) nested within lab sections. The different models that were run are shown in Supplemental Table S1. The first model only included the covariates (gender, ethnicity, low-income status, first-generation status, SAT or ACT composite scaled score). This first model showed that these variables alone (except for the low-income covariate) were significantly related to grades (which is consistent with prior research), and the R2 value with only these variables was 0.05. The ethnicity variable was dummy coded (Asian, Black, Hispanic, White) and only three of the four dummy variables (Asian, Black, Hispanic) were included (White was the reference group). The second model only included the match-gender variable of interest, and this variable alone (without any covariates) was unrelated to grades. When we included the covariates and this variable of interest (model 3), the R^2 value was 0.05, because it included the covariates. We replicated this for the other variables of

interest (match ethnicity, match low income, match first generation) in models 4–9.

All models accounted for this nested structure of the data, with undergraduates clustered within the same lab section with the same GTA using a cluster command in Stata (v. 16). The decision to cluster on lab section rather than quarter or year was based both on substantive and statistical considerations. For example, we compared grades by year and found initial differences, $F(3, 972) = 6.77$, $p < 0.001$. Post hoc tests indicated that, although students entering in the Fall 2014 (M = 2.36, SD = 0.92), Fall 2015 (M = 2.37, SD = 0.84), and Fall 2016 (M = 2.31, SD = 0.91) had similar grades, students entering in Fall 2017 had higher grades (M = 2.64, SD = 0.85). To investigate this further, we ran regression analyses separately for each year and combined all of the years and included a dummy variable for year. The results were the same, so we decided not to cluster on year. We conducted a similar investigation to check for differences by quarter. Instead of clustering on year or quarter, we decided to cluster on lab section, because this reflects unique experiences for students who were in a particular lab section for a particular quarter and year. This clustering reflects substantive similarities in the experiences of students in the same lab sections with a particular GTA. In addition to determining how to cluster the data, we also investigated other modeling methods, such as hierarchical linear modeling (e.g., see Raudenbush and Bryk, 2001). Given similar results, what is reported are the most parsimonious models.

RESULTS

More than half of the undergraduate students and GTAs matched in terms of gender (51%), but only 20% matched in terms of race/ethnicity. Matching for low-income and first-generation status was 72 and 73%, respectively. Table 3 suggests that performance in the biology lab course was higher for students who matched their GTAs in terms of ethnicity, $F(1, 974) = 4.64$, $p < 0.05$, and first-generation status, $F(1, 974) = 4.06$, $p < 0.05$, compared with students who did not match in terms of race/ethnicity and first-generation status. There were no differences in performance based on a gender match, $F(1, 974) = 0.78$, $p = 0.38$, or a low-income match, $F(1, 974) = 0.78$, $p = 0.06$, between students and their GTAs. In other words, students who matched with their GTAs in terms of gender performed similarly to students who did not match with their GTAs in terms of gender. While not statistically different ($p = 0.06$), there was a trend for students who matched with their GTAs in terms of low-income status to have higher grades compared with students who did not match.

The next step in the analyses used regression analyses to compare academic performance that included the matching variables (Table 3) and covariates (demographics, prior academic performance) and that accounted for the nested structure of the data. Although initial descriptive information suggested no differences in course grades for matching in terms of gender and low-income status; we ran regression models with covariates to statistically test for all four matching characteristics (race/ethnicity, gender, low-income status, and first-generation status; full results are provided in Supplemental Table S1). Results from these separate analyses suggest that matching demographics of undergraduates and GTAs was not predictive of undergraduate students' performance in a biology lab course

TABLE 3. Descriptive statistics for course grades by matching variables ($n = 976$)

	Match			Not matched			
	n	M	SE	n	M	SE	$F(1, 974)$
Gender	508	2.44	0.04	468	2.40	0.04	0.78
Ethnicity	281	2.52	0.05	695	2.39	0.03	4.64*
Low income	678	2.46	0.03	298	2.34	0.05	3.62
First generation	704	2.46	0.03	272	2.33	0.05	4.06*

*$p < 0.05$.

(Table 4). The nonsignificant predictors and similar R^2 values across all models indicate that these matching variables did not explain the variation in biology grades.

To further explore which students had higher grades, we focused on students who matched with their GTAs in terms of different characteristics. Of those students who matched with their GTAs in terms of gender ($n = 508$), there were differences in course grades, $F(1, 506) = 8.35$, $p < 0.01$. Post hoc comparisons indicated that male students with a male GTA (M = 2.54, SE = 0.05, $n = 318$) performed higher than female students with a female GTA (M = 2.30, SE = 0.06, $n = 190$; Tukey HSD = 4.09, $p < 0.05$; Figure 3A). There were also differences in course grades for those students who matched in terms of race/ethnicity ($n = 281$), $F(2, 277) = 7.61$, $p < 0.001$ (Figure 3B). White students with White GTAs (M = 2.87, SE = 0.09, $n = 66$) performed higher than Asian students with Asian GTAs (M = 2.44, SE = 0.06, $n = 193$; Tukey HSD = 3.46, $p < 0.05$) and Latinx students with Latinx GTAs (M = 2.27, SE = 0.17, $n = 22$; Tukey HSD = 4.87, $p < 0.05$). There were no differences in grades for students who matched with their GTAs in terms of low-income status, $F(1, 676) = 0.01$, $p = 0.93$ (Figure 3C), or who matched with their GTAs in terms of first-generation status, $F(1, 702) = 1.29$, $p = 0.26$ (Figure 3D).

Female Student Performance when Matched with Female GTAs

In this analysis, the models included all students. However, course grades for students who matched with their GTAs in terms of gender and ethnicity were different. To further explore these findings for subgroups of students who are typically underrepresented in STEM fields, we analyzed female and Latinx students' academic performances. As noted earlier, overall, female students performed lower than male students; however, we explored the potential added benefit or effect of having a female GTA. For female students, there was no significant difference in performance for female students who had a female GTA (M = 2.30, SE = 0.06) or male GTA (M = 2.34, SE = 0.05, $t(534) = 0.47$, $p = 0.64$). In other words, a match in gender between female students and their GTAs did not relate to higher performance in the introductory biology course.

Latinx Student Performance when Matched with Latinx GTAs

There was no difference in performance in the introductory biology course for Latinx students if they had a Latinx GTA, M = 2.27, SE = 0.17, or non-Latinx GTA, M = 2.34, SE = 0.05, $t(318) = 0.35$, $p = 0.73$. Thus, similar to females, Latinx students did not have higher or lower biology course grades when they had a Latinx GTA. To explore the impact further for Latinx students, we focused specifically on the intersection of gender and race/

ethnicity. There was no significant difference in performance for Latina students who had a GTA with same gender, $t(201) = 0.65$, $p = 0.52$, or ethnicity, $t(201) = -0.04$, $p = 0.97$.

DISCUSSION

Although this study occurred in a structurally diverse institution, similar to STEM programs across the nation, the demographics of the sample of GTAs included in this study differed from the undergraduate population. In this study, URM students represented almost 35% of the undergraduate students but only 14% of the GTA population. This is consistent with other studies that have reported Latinx undergraduate student enrollment is higher than Latinx graduate student enrollment at HSIs (Garcia, 2017; Garcia and Guzman-Alvarez, 2019). Similarly, the number of URM students earning bachelor's degrees in STEM is greater than the number earning doctoral degrees in STEM (Estrada *et al.*, 2016).

While increasing representation and success in advanced STEM fields is an ongoing issue, one goal of this study was to determine whether there was an impact on academic performance for students where there was a match between the demographics of the undergraduates and GTAs. For this particular sample, there were no significant relationships between matching and performance in an introductory biology laboratory course. This finding is consistent with a study that concluded GTA demographics were unrelated to students' content knowledge of chemistry in a population of predominantly White GTAs and White undergraduate students (Wheeler *et al.*, 2017). The instructor-matching literature is limited in higher education, and our study adds to this body of work, looking specifically at GTA instructor level and underrepresented students at an HSI.

Matching the gender and/or race/ethnicity of the GTA did not translate to improved academic performance for female and URM students. It is possible that the effects of instructor matching may not directly influence the course grade. Instead, it may positively affect other student outcomes, including female and URM students' confidence and persistence in STEM (Cotner *et al.*, 2011). Interestingly, when female students are taught by a female GTA and a male faculty instructor, female students exhibit increased confidence in critical thinking and discussing scientific findings, but not in content knowledge (Cotner *et al.*, 2011). Students in this study concurrently enrolled in the lab course along with a lecture course. In the lecture course, students were taught by mostly male faculty (65%), which may have affected their confidence in content knowledge. Another possible positive outcome of GTA instructor matching may be related to persistence in STEM. With more female and URM graduate students, there is an increased likelihood of persistence for female and URM undergraduate
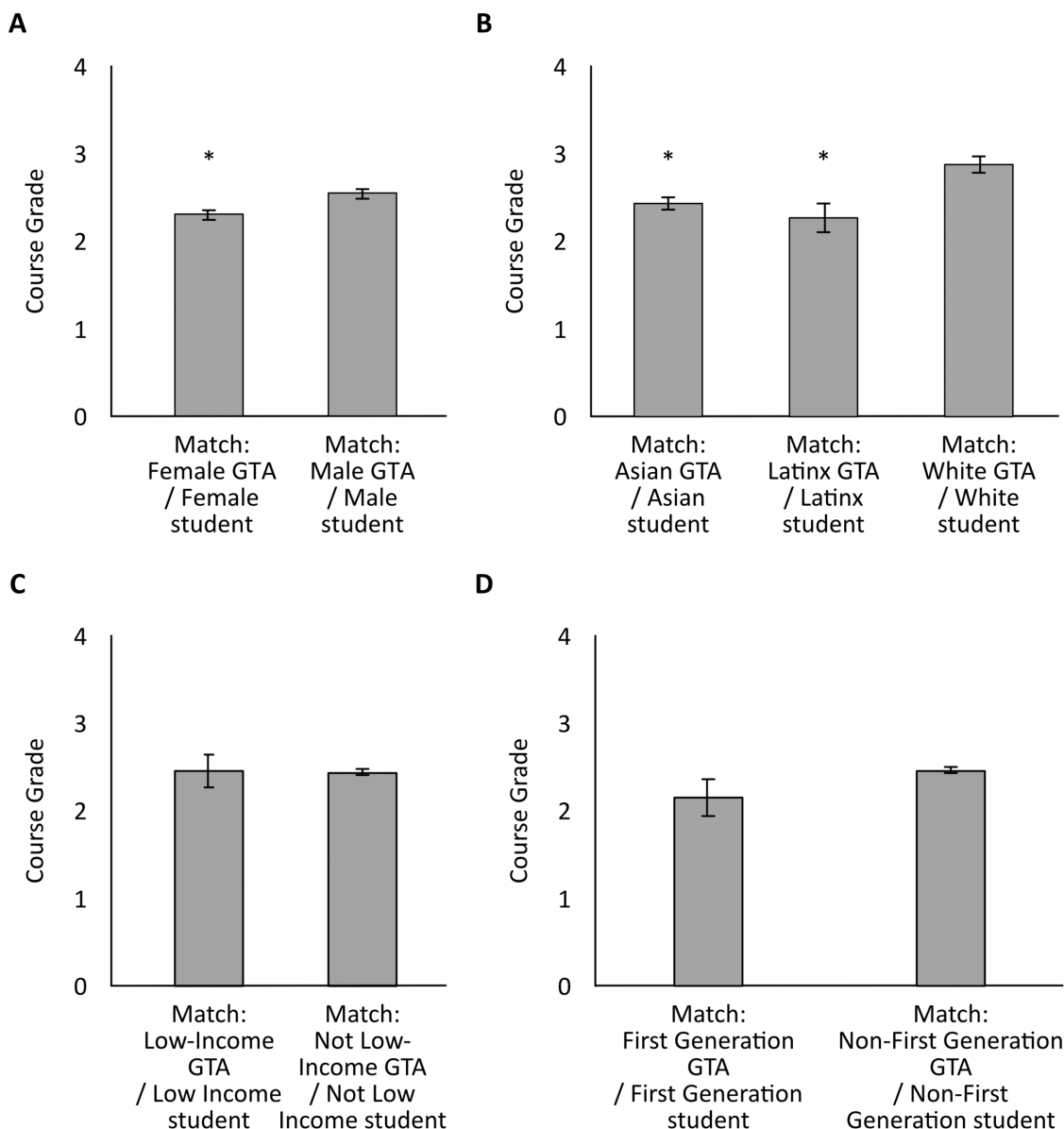
FIGURE 3. Comparison of course grades for students who matched with their GTAs in terms of gender (A, *n* = 508), ethnicity (B, *n* = 281), low-income status (C, *n* = 678), and first-generation status (D, *n* = 704). (A) Differences in course grades for students based on matching gender with their GTA, *F*(1, 506) = 8.35, *p* < 0.01. Of students who matched with their GTAs, male students taught by male GTAs (M = 2.54, SE = 0.05, *n* = 318) received higher grades than female students taught by female GTAs (M = 2.30, SE = 0.05, *n* = 190; Tukey HSD = 4.09, *p* < 0.05). (B) Differences in course grades for students who matched with their GTAs in terms of race/ethnicity, *F*(2, 277) = 7.61, *p* < 0.001. Of the students who matched with their GTAs, White students taught by White GTAs (M = 2.87, SE = 0.09, *n* = 66) received higher grades than both Asian students taught by Asian GTAs (M = 2.44, SE = 0.06, *n* = 193; Tukey HSD = 3.46, *p* < 0.05) and Latinx students taught by Latinx GTAs (M = 2.27, SE = 0.17, *n* = 22; Tukey HSD = 4.87, *p* < 0.05). There were no differences in grades for students who matched with their GTAs in terms of low-income status (C), *F*(1, 676) = 0.01, *p* = 0.93), or first-generation status (D) *F*(1, 702) = 1.29, *p* = 0.26). All data are presented as mean ± SE.

students in STEM (Griffith, 2010). Future research could explore the effects of GTA instructor matching on other student outcomes, including persistence and confidence of female and URM students.

Instructor matching has been shown to be effective in elementary and secondary education based on the teaching beliefs and practices of teachers of color (Villegas and Irvine, 2010). These studies suggest that teachers of color, in addition

to sharing gender and/or race/ethnicity with students, typically have high expectations for all their students, develop trust with their students, and practice culturally relevant/responsive teaching (CRP; Villegas and Irvine, 2010). The teaching practices of teachers of color contribute to the success of their students of color. URM students do better when instructors implement student-centered teaching practices, such as active-learning strategies, compared with teacher-centered

TABLE 4. Summary multiple regression results (n = 976)[a]

	Model 1		Model 2		Model 3		Model 4	
	B	SE B	B	SE B	B	SE B	B	SE B
Match–Gender	−0.03	0.07						
Match–Ethnicity			0.05	0.07				
Match–Low Income					0.03	0.07		
Match–First Generation							−0.14	0.10
$F_{(8, 119)}$	8.45*		8.67*		8.66*		8.56*	
R^2	0.05		0.05		0.05		0.05	

[a]All models include student demographics (gender, ethnicity, first-generation status, low-income status) and prior achievement (SAT/ACT score).
*$p < 0.001$.

practices (Haak *et al.*, 2011; Eddy and Hogan, 2014). Implementing student-centered teaching practices and CRP requires significant professional development (Morrison *et al.*, 2008; Henderson *et al.*, 2011; Johnson, 2011). The professional development provided to the GTAs in this study was minimal, and GTAs were not exposed to CRP. Given their limited amount of pedagogy training, it is likely the GTAs' teaching beliefs and practices were mainly teacher focused (Gormally *et al.*, 2016; Lee, 2019). The benefits of instructor matching for female and URM students may be more salient when instructors practice student-centered teaching and CRP.

Limitations
Racial/ethnic groups are often categorized using broad categories: White, Black, Asian, and Latinx. Within each racial/ethnic category, many subgroups can be identified (Perez and Hirschman, 2009). For example, Black may include students who are African American, Jamaican, Haitian, and Nigerian, among many others. Asian may include students who are Chinese, Filipinx, Indian, Cambodian, and Hmong, among many others. Latinx may include students who are Mexican, Puerto Rican, Salvadoran, Colombian, and Venezuelan, among many others. Studies have shown that student enrollment, classroom engagement, and academic performance within the subgroups are different (Nuñez and Kim, 2012; Cuellar, 2014; Ing and Victorino, 2016). Additionally, given the cultural and identity differences, it would be more accurate to match students and GTAs along the subgroups to determine whether matching demographics have an impact on student performance.

In this particular GTA sample, there was only one GTA who identified as Black, who was excluded from the analysis. Additionally, there were only 31 Black undergraduate students included in the analysis, all of whom were taught by a non-Black GTA. Based on the limited number of Black GTAs and undergraduate students, it was not possible to determine whether GTA instructor matching had an impact on academic performance for Black undergraduates. This highlights an ongoing problem in the recruitment and retention of Black students in STEM.

There were six Latinx GTAs included in this study. Even at an HSI, it is unlikely that Latinx students will be taught by Latinx GTAs in an introductory biology laboratory course. The sample size was small for Latinx students who matched with Latinx GTAs. Given the small sample size, we were unable to thoroughly investigate the impact of GTA instructor matching both gender and race/ethnicity on Latina and Latino students. The impact of matching remains unknown for Latina students taught by Latina GTAs and Latino students taught by Latino GTAs.

There were no GTAs and less than 0.5% undergraduate students who identified as nonbinary in the study, so we could not determine whether GTA instructor matching for nonbinary students had an impact on academic performance. The institution does not collect data on gender-nonconforming or transgender students. It is unknown whether GTAs would share their gender identity with their students. Similarly, it is unknown whether GTAs share their income status or being first generation with their students. The impact of matching of gender identity and low-income and/or first-generation status may depend on GTAs sharing the information with students. It is also possible that GTAs from nontraditional backgrounds may be more open and strategic in creating inclusive classrooms. There is a lack of research focused on the teaching beliefs and practices of GTAs from nontraditional backgrounds.

Call to Action
This study evaluated the impact of GTA instructor matching in first-year students majoring in the life sciences at an HSI. Results suggest no improvement attributable to matching in students' academic performance in an introductory biology lab course. More specifically, female and Latinx students did not improve their course grades when taught by gender- and/or race/ethnicity-congruent GTAs. As mentioned earlier, the effects of GTA matching on Black, gender-nonconforming, non-binary, and transgender students remains unknown. However, to further evaluate the effects of GTA instructor matching, future research should include a larger sample of underrepresented GTAs across race/ethnicity and the gender spectrum.

In addition to a need for a more diverse graduate student population, continued attention to the recruitment and retention of underrepresented students is critical. While researchers have called for moving beyond structural diversity (Denson and Chang, 2009), the "chilly" STEM climate that impacts underrepresented students needs to be eliminated (Seymour and Hewitt, 1997), but this is not as simple as just putting more underrepresented GTAs in the classroom or increasing the diversity of the undergraduate population (Hurtado *et al.*, 2010; Ong *et al.*, 2018). We need to improve how we teach students at all levels, particularly GTAs who teach a large number of STEM undergraduates. Though it is important to keep in mind that pedagogical strategies may affect students differentially (Aguillon *et al.*, 2020), incorporating CRP and student-centered strategies into GTA professional development may benefit underrepresented students. Regardless of their demographic

characteristics, it is important to develop and cultivate self-aware and empathetic GTAs who are better equipped to implement pedagogical strategies and improve the classroom climate (Dewsbury and Brame, 2019; Dewsbury, 2020). While female and URM GTAs may already be acting as mentors and role models for female and URM undergraduate students (Ong *et al.*, 2011; Marin and Pereschica, 2017), the responsibility of creating inclusive classrooms should not be shouldered by only underrepresented GTAs. All GTAs can learn how to improve their teaching practices to positively impact on the student outcomes of students typically underrepresented in STEM fields.

REFERENCES

Aguillon, S. M., Siegmund, G.-F., Petipas, R. H., Drake, A. G., Cotner, S., & Ballen, C. J. (2020). Gender differences in student participation in an active-learning classroom. *CBE—Life Sciences Education*, *19*(2), ar12.

Alcantar, C. M., & Hernandez, E. (2020). "Here the professors are your guide, tus guías": Latina/o student validating experiencing with faculty at a Hispanic-serving community college. *Journal of Hispanic Higher Education*, *19*(1), 3–18. https://doi.org/10.1177/1538192718766234.

Bailey, E. G., Greenall, R. F., Baek, D. M., Morris, C., Nelson, N., Quirante, T. M., … & Williams, K. R. (2020). Female in-class participation and performance increase with more female peers and/or a female instructor in life sciences courses. *CBE—Life Sciences Education*, *19*(3), ar30.

Bettinger, E. P., & Long, B. T. (2005). Do faculty serve as role models? The impact of instructor gender on female students. *American Economic Review*, *95*(2), 152–157.

Blake-Beard, S., Bayne, M. L., Crosby, F. J., & Muller, C. B. (2011). Matching by race and gender in mentoring relationships: Keeping our eyes on the prize. *Journal of Social Issues*, *67*(3), 622–643.

Carlone, H. B., & Johnson, A. (2007). Understanding the science experiences of successful women of color: Science identity as an analytic lens. *Journal of Research in Science Teaching*, *44*(8), 1187–1218.

Carrell, S. E., Page, M. E., & West, J. E. (2010). Sex and science: How professor gender perpetuates the gender gap. *Quarterly Journal of Economics*, *125*(3), 1101–1144.

Clark, S. L., Dyar, C., Maung, N., & London, B. (2016). Psychosocial pathways to STEM engagement among graduate students in the life sciences. *CBE—Life Sciences Education*, *15*(3), ar45.

Cotner, S., Ballen, C., Brooks, D. C., & Moore, R. (2011). Instructor gender and student confidence in the sciences: A need for more role models? *Journal of College Science Teaching*, *40*(5), 96–101.

Cuellar, M. (2014). The impact of Hispanic-Serving Institutions (HSIs), emerging HSIs, and Non-HSIs on Latina/o academic self-concept. *Review of Higher Education*, *37*(4), 499–530.

Denson, N., & Chang, M. J. (2009). Racial diversity matters: The impact of diversity-related student engagement and institutional context. *American Educational Research Journal*, *46*(2), 322–353.

Dewsbury, B. M. (2020). Deep teaching in a college STEM classroom. *Cultural Studies of Science Education*, *15*, 169–191.

Dewsbury, B. M., & Brame, C. J. (2019). Inclusive teaching. *CBE—Life Sciences Education*, *18*(2), fe2.

Eagan, K., Hurtado, S., Figueroa, T., & Hughes, B. (2014). *Examining STEM pathways among students who begin college at four-year institutions (Paper commissioned for the Committee on Barriers and Opportunities in Completing 2- and 4-Year STEM Degrees)*. Washington, DC: Retrieved July 12, 2019, from http://sites.nationalacademies.org/cs/groups/dbassesite/documents/webpage/dbasse_088834.pdf

Eddy, S. L., Brownell, S. E., & Wenderoth, M. P. (2014). Gender gaps in achievement and participation in multiple introductory biology classrooms. *CBE—Life Sciences Education*, *13*(3), 478–492.

Eddy, S. L., & Hogan, K. A. (2014). Getting under the hood: How and for whom does increasing course structure work? *CBE—Life Sciences Education*, *13*(3), 453–468.

Estrada, M., Burnett, M., Campbell, A. G., Campbell, P. B., Denetclaw, W. F., Gutiérrez, C. G., … & Zavala, M. (2016). Improving underrepresented minority student persistence in STEM. *CBE—Life Sciences Education*, *15*(3), es5.

Fairlie, R. W., Hoffmann, F., & Oreopoulos, P. (2014). A community college instructor like me: Race and ethnicity interactions in the classroom. *American Economic Review*, *104*(8), 2567–2591.

Garcia, G. A. (2017). What does it mean to be Latinx-serving? Testing the utility of the typology of HSI organizational identities. *Association of Mexican American Educators Journal*, *11*(3), 109–138.

Garcia, G. A., & Guzman-Alvarez, A. (2019). Descriptive analysis of graduate enrollment trends at Hispanic-serving institutions: 2005-2015. *Journal of Hispanic Higher Education*, https://doi.org/10.1177/1538192719835681.

Good, C., Rattan, A., & Dweck, C. S. (2012). Why do women opt out? Sense of belonging and women's representation in mathematics. *Journal of Personality and Social Psychology*, *102*(4), 700–717.

Gormally, C., Sullivan, C. S., & Szeinbaum, N. (2016). Uncovering barriers to teaching assistants (TAs) implementing inquiry teaching: Inconsistent facilitation techniques, student resistance, and reluctance to share control over learning with students. *Journal of Microbiology & Biology Education*, *17*(2), 215–224.

Griffith, A. L. (2010). Persistence of women and minorities in STEM field majors: Is it the school that matters? *Economics of Education Review*, *29*(6), 911–922.

Haak, D. C., HilleRisLambers, J., Pitre, E., & Freeman, S. (2011). Increased structure and active learning reduce the achievement gap in introductory biology. *Science*, *332*(6034), 1213–1216.

Hagedorn, L. S., Chi, W., Cepeda, R., & McLain, M. (2007). An investigation of critical mass: The role of Latino representation in the success of urban community college students. *Research in Higher Education*, *48*(1), 73–91.

Hartman, H., & Hartman, M. (2008). How undergraduate engineering students perceive women's (and men's) problems in science, math, and engineering. *Sex Roles*, *58*(3–4), 251–265.

Henderson, C., Beach, A., & Finkelstein, N. (2011). Facilitating change in undergraduate STEM instructional practices: An analytical review of the literature. *Journal of Research in Science Teaching*, *48*(8), 952–984.

Hoffmann, F., & Oreopoulos, P. (2009). A professor like me: The influence of instructor gender on college achievement. *Journal of Human Resources*, *44*(2), 479–494.

Hurtado, S., & Carter, D. F. (1997). Effects of college transition and perceptions of the campus racial climate on Latino college students' sense of belonging. *Sociology of Education*, *70*(4), 324–345.

Hurtado, S., Eagan, M. K., Tran, M. C., Newman, C. B., Chang, M. J., & Velasco, P. (2011). "We do science here": Underrepresented students' interactions with faculty in different college contexts. *Journal of Social Issues*, *67*(3), 553–579.

Hurtado, S., Newman, C. B., Tran, M. C., & Chang, M. J. (2010). Improving the rate of success for underrepresented racial minorities in STEM fields: Insights from a national project. *New Directions for Institutional Research*, *148*, 5–15.

Ing, M., & Victorino, C. (2016). Differences in classroom engagement of Asian American engineering students. *Journal of Engineering Education*, *105*(3), 431–451.

Johnson, C. C. (2011). The road to culturally relevant science: Exploring how teachers navigate change in pedagogy. *Journal of Research in Science Teaching*, *48*(2), 170–198.

Johnson, D. R. (2012). Campus racial climate perceptions and overall sense of belonging among racially diverse women in STEM majors. *Journal of College Student Development*, *53*(2), 336–346.

Lee, S. W. (2019). The impact of a pedagogy course on the teaching beliefs of inexperienced graduate teaching assistants. *CBE—Life Sciences Education*, *18*(1), ar5.

Lundberg, C. A., & Schreiner, L. A. (2004). Quality and frequency of faculty-student interaction as predictors of learning: An analysis by student race/ethnicity. *Journal of College Student Development*, *45*(5), 549–565.

Maestas, R., Vaquera, G. S., & Zehr, L. M. (2007). Factors impacting sense of belonging at a Hispanic-serving institution. *Journal of Hispanic Higher Education*, *6*(3), 237–256.

Marin, P., & Pereschica, P. (2017). Becoming an Hispanic-serving research institution: Involving graduate students in organizational change. *Association of Mexican American Educators Journal*, *11*(3), 154–175.

Morrison, K. A., Robbins, H. H., & Rose, D. G. (2008). Operationalizing culturally relevant pedagogy: A synthesis of classroom-based research. *Equity & Excellence in Education*, *41*(4), 433–542.

National Academy of Sciences, National Academy of Engineering, and Institute of Medicine. (2011). *Expanding underrepresented minority participation: America's science and technology talent at the crossroads*. Washington, DC: National Academies Press. https://doi.org/10.17226/12984.

National Academies of Sciences, Engineering, and Medicine. (2019). *Minority serving institutions: America's underutilized resource for strengthening the STEM workforce*. Washington, DC: National Academies Press. https://doi.org/10.17226/25257.

National Science Foundation, National Center for Science and Engineering Statistics. (2019). *Women, minorities, and persons with disabilities in science and engineering: 2019 (Special Report NSF 19-304)*. Alexandria, VA:.

Nelson Laird, T. F., Bridges, B. K., Morelon-Quainoo, C. L., Williams, J. M., & Holmes, M. S. (2007). African American and Hispanic student engagement at minority serving and predominantly White institutions. *Journal of College Student Development*, *48*(1), 39–56.

Nuñez, A. M., & Kim, D. (2012). Building a multicontextual model of Latino college enrollment: Student, school, and state-level effects. *Review of Higher Education*, *35*(2), 237–263.

Ong, M., Smith, J. M., & Ko, L. T. (2018). Counterspaces for women of color in STEM higher education: Marginal and central spaces for persistence and success. *Journal of Research in Science Teaching*, *55*(2), 206–245.

Ong, M., Wright, C., Espinosa, L. L., & Orfield, G. (2011). Inside the double bind: A synthesis of empirical research on undergraduate and graduate women of color in science, technology, engineering, and mathematics. *Harvard Educational Review*, *81*(2), 172–208.

Perez, A. D., & Hirschman, C. (2009). The changing racial and ethnic composition of the US population: Emerging American identities. *Population and Development Review*, *35*(1), 1–51.

Raudenbush, S., & Bryk, R. (2001). *Hierarchical linear models: Applications and data analysis methods*. Newbury Park, CA: Sage.

Robst, J., Keil, J., & Russo, D. (1998). The effect of gender composition of faculty on student retention. *Economics of Education Review*, *17*(4), 429–439.

Rushin, J. W., DeSaix, J., Lumsden, A., Streubel, D. P., Summers, G., & Bernson, C. (1997). Graduate teaching assistant training: A basis for improvement of college biology teaching & faculty development? *American Biology Teacher*, *59*(2), 86–90.

Seymour, E., & Hewitt, N. M. (1997). *Talking about leaving: Why undergraduates leave the sciences*. Boulder, CO: Westview Press.

Sundberg, M. D., Armstrong, J. E., & Wischusen, W. (2005). A reappraisal of the status of introductory biology laboratory education in U.S. colleges & universities. *American Biology Teacher*, *67*(9), 525–529.

Villegas, A. M., & Irvine, J. J. (2010). Diversifying the teaching force: An examination of major arguments. *Urban Review: Issues and Ideas in Public Education*, *42*(3), 175–192.

Walton, G. M., & Cohen, G. L. (2011). A brief social-belonging intervention improves academic and health outcomes of minority students. *Science*, *331*(6023), 1447–1451.

Wheeler, L. B., Maeng, J. L., Chiu, J. L., & Bell, R. L. (2017). Do teaching assistants matter? Investigating relationships between teaching assistants and student outcomes in undergraduate science laboratory classes. *Journal of Research in Science Education*, *54*(4), 463–492.