Participation in Undergraduate Research Reduces Equity Gaps in STEM Graduation Rates

Heather Haeger,[†]* Elia Hilda Bueno,[‡] and Quentin Sedlacek[§]
[†]STEM Learning Center, Educational Policy Studies and Practices, and [‡]Human Development and Family Science, University of Arizona, Tucson, AZ 85721; [§]Department of Teaching & Learning, Simmons School of Education & Human Development, Southern Methodist University, Dallas, TX 75275

ABSTRACT
Many students who enroll in a public U.S. 4-y college will not graduate. The odds of completing a college degree are even lower for students who have been marginalized in higher education, especially in Science, Technology, Engineering, and Math (STEM) fields. Can undergraduate research increase a student's likelihood of graduating college and close educational equity gaps in college completion? To answer this question, we use data from six public U.S. universities (*N* = 120,308 students) and use Propensity Score Matching to generate a comparison group for analyses. We conducted logistic regressions on graduation rates and equity gaps in 4 and 6 y using the matched comparison group and undergraduate researchers in STEM (*n* = 2727). When being compared with like-peers and controlling for background characteristics and prior academic performance, students who participated in undergraduate research were twice as likely to graduate in 4 y and over 10 times as likely to graduate in 6 y. We also found that equity gaps in 4-y graduation rates for students of color, low-income, and first-generation students were cut in half for undergraduate researchers. At 6 y, these gaps were completely closed for undergraduate researchers. As we seek ways to close education gaps and increase graduation rates, undergraduate research can be a meaningful practice to improve student success.

INTRODUCTION
High-impact practices such as undergraduate research (UR) have demonstrated a compensatory effect where students who have been marginalized in education because of their socioeconomic status or their race/ethnicity benefit the most from these practices (Kinzie *et al.* 2008). Though the compensatory effect of UR has not been systematically tested on graduation rates, prior studies substantiate why research participation may increase graduation rates and reduce equity gaps. The opportunity for students to work with faculty on research can be a transformative experience (Olivares-Donoso and González, 2019). Students who participate in UR build important connections within their major (Aikens *et al.*, 2016), develop resiliency and problem-solving strategies (Cartrette and Melroe-Lehrman 2012), and feel more prepared for graduate education (Hathaway *et al.*, 2002). Additionally, students who participate in research identify more with their field (Hurtado *et al.*, 2011), build critical skills like communication and teamwork (Carter *et al.*, 2016), and report greater self-efficacy and higher-order learning (Kuh *et al.*, 2010; Robnett *et al.*, 2015).

Exposure to UR may also enhance STEM persistence among marginalized student populations. Students of color are more likely to encounter academic obstacles including lower levels of sense of belonging (Hausmann *et al.*, 2009) and STEM identity (Chang *et al.*, 2011) compared with their White and Asian peers. Different external factors can hinder persistence for marginalized students including social factors that impact first-generation college students (Stephens *et al.*, 2012), larger institutional

Erin Shortlidge, *Monitoring Editor*
Submitted Mar 23, 2022; Revised Dec 5, 2023; Accepted Jan 4, 2024
CBE Life Sci Educ March 1, 2024 23:ar11
DOI:10.1187/cbe.22-03-0061
*Address correspondence to: Heather Haeger (heather3@arizona.edu).

structures and norms that reinforce racial bias (Chang *et al.*, 2011), and individual discriminatory experiences based on social class (Yee, 2016). While marginalized students possess cultural strengths (e.g., in-home knowledge, life experiences, and resourcefulness) these assets are often undervalued by higher education institutions (Kiyama and Rios-Aguilar, 2017). Similarly, hostile STEM interactions (i.e., unsupportive professors and peers) and lack of relatability with course material can discourage overall degree completion (Johnson, 2012; Smith and Lucena, 2016; Wilkins-Yel *et al.*, 2022). Thus, it is necessary to investigate interventions such as UR that enhance interest and success in STEM, increase representation in academic STEM settings, and diversify the STEM workforce. In the present study, we consider UR as a potential mechanism for increasing equity by first examining some of the sources of the equity gaps in college completion, then discussing how UR might mitigate these inequities, and finally by conducting a post hoc quasiexperimental study on the relationship between UR and college graduation rates.

Sources of the Equity Gap in College Completion

Many students who pursue a college degree never finish their degree. An incoming student at a public U.S. 4-y college only has a 36% chance of graduating in 4 y and a 59% chance of graduating in 6 y (NCES, 2019). The odds of completing a college degree are even lower for students who have been traditionally marginalized and underserved in higher education because of their socioeconomic status or their race/ethnicity, especially in Science, Technology, Engineering, and Math (STEM) fields. Low-income and first-generation college students are less likely to graduate college (Wilbur and Roscigno, 2016) but those that do graduate receive tremendous economic benefits over their lifetime (Stephens *et al.*, 2014; Manzoni and Streib, 2019). Similarly, students of color in STEM are more likely to transfer out STEM or stop out of college completely (NCES, 2019; Riegle-Crumb *et al.*, 2019). These equity gaps for marginalized students in STEM fields have dramatic economic and social ramifications. Individuals who earn a bachelor's degree have 84% higher lifetime earnings compared with those who have a high school diploma and 50% higher lifetime earnings than those who started college but never earned a bachelor's degree (Carnevale *et al.*, 2011). Effects of racial discrimination, gender discrimination, and other forms of inequality may compound and amplify these disparities over the course of individuals' lives (Nunley *et al.*, 2015).

Equity gaps between traditionally privileged students (e.g., higher-SES students, White students, and students whose parents have college degrees) and marginalized students (e.g., lower-income students, students of color, and first-generation college students) have numerous structural and systemic causes. A full accounting of these causes is beyond the scope of this paper. Instead, we focus on specific mechanisms of economic, social, and cultural capital by which some students are disadvantaged in college while others are comparatively privileged (Bourdieu, 1986; Walpole, 2003).

Disparities in economic capital can affect college completion as students from low-income families may struggle to afford college tuition and associated costs (Cabrera *et al.*, 1990; Goldrick-Rab, 2016). They may also struggle with food or housing insecurity and the physical and emotional stresses that accompany these conditions (Broton and Goldrick-Rab, 2018; Goldrick-Rab *et al.*, 2019). Many low-income students meet their expenses through part- or full-time employment during school; however, employment that surpasses 15 h per week is associated with reduced chances of college completion (Perna, 2010). This effect may happen because of conflicts between the time required for employment and the time required for coursework.

Even when school and work schedules do not conflict, students with jobs may be unable to participate in social gatherings or extracurricular opportunities which could otherwise enable them to form stronger relationships with college faculty, staff, and peers (Means and Pyne, 2017). Thus, a lack of economic capital may prevent students from accessing specific forms of social capital that could facilitate their degree completion and long-term success. We should note that students employed during college often develop invaluable knowledge, experience, and social relationships through their employment; however, higher education institutions do not always value these relationships, experiences, and knowledge (McClellan *et al.*, 2018).

Disparities in cultural capital – for example, disparities in parental education – may also affect college completion. All students and all parents possess a wealth of valuable knowledge, skills, perspectives, and experiences. However, institutions of higher education are designed to value specific (and often narrowly defined) knowledge, skills, perspectives, and experiences (Yosso, 2005). Students whose parents did not complete 4-y degrees may be less knowledgeable about the hidden curricula of academia compared with higher-SES peers whose parents did complete such degrees (Cabrera and La Nasa, 2000). Furthermore, they may perceive themselves as "not belonging" within the university environment. This feeling of "belonging uncertainty" can make students reluctant to take advantage of university resources that might help them develop the specific forms of cultural capital valued by the university (Walton and Cohen, 2007, 2011). Thus, students who lack a specific and narrowly defined form of cultural capital (i.e., parental knowledge of college) may be less likely than their more privileged peers to access university resources and experiences such as UR that could facilitate their degree completion and long-term success.

These disparities in economic, social, and cultural capital can contribute to inequities in college completion. However, these inequities are not inevitable. At their inception, higher education institutions in the United States primarily served a narrow sector of the population: relatively high-SES, male, European American students (Thelin, 2011). With the exception of Minority Serving Institutions such as Historically Black Colleges and Universities or Tribal Colleges, institutions were designed with students from this narrow sector of the population in mind (*Gasman et al.*, 2015). However, in recent decades, many institutions have sought to redesign themselves in order to better serve the diverse population of the United States. One way in which institutions have sought to serve their diverse student bodies has been through UR (Jones *et al.*, 2010; Eagan *et al.*, 2013). Forms of capital influence who participates in UR through economic barriers, access to social capital related to research, and cultural capital in college for the students who stand to benefit the most from participation (Kinzie *et al.*, 2008). UR has shown promise at reducing inequities in student outcomes. We argue that it may support equity through its impacts on economic capital related to future earnings from degree completion, social capital through faculty and peer connections, and cultural capital in increased knowledge of STEM career options.

UR: A Mechanism to Reduce Equity Gaps

UR experiences may disrupt mechanisms of inequality by promoting more equitable student access to specific forms of economic, social, and cultural capital (Aikens *et al.*, 2016; Thompson and Jensen-Ryan, 2018). For example, marginalized students may not have the time or financial resources to engage in unpaid research work – but if these students have access to *paid* research opportunities, UR could help students add to or maintain their personal finances rather than impeding them from participation. In this way, UR can provide students with economic capital while simultaneously helping students develop specific forms of social and cultural capital that are traditionally valued by universities. Other types of student employment (e.g., working in a campus bookstore or restaurant) may help students develop valuable economic, social, and cultural capital as well; however, these forms of social and cultural capital are less likely to include the specific forms traditionally valued by universities.

UR can support students' development of social capital in several ways. For example, students who participate in mentored UR may have transformative experiences collaborating with university faculty (Olivares-Donoso and González, 2019) and even more positive experiences working with both faculty and graduate students or postdoctoral researchers (Aikens *et al.*, 2016). These social relationships can help students develop strong identities as researchers and may alleviate students' uncertainty about their sense of belonging within the institution or the discipline, leading them to respond to challenges and adversity in adaptive ways (Robnett *et al.*, 2015). Furthermore, these social relationships can cultivate students' underlying motivations for learning different disciplinary knowledge or skills. For example, ethnographic research by Artemeva (2011) illustrated how social relationships with research team members can lead an undergraduate to engage more deeply and reflectively in important activities such as science writing. For students who work in courses, labs, or research teams with fellow undergraduates, group activities with peers can also contribute to interpersonal relationships that support students' sense of self-efficacy, short-term learning, and long-term academic success (Cohen *et al.*, 1999).

Research experiences can also directly or indirectly help students develop numerous forms of cultural capital valued by the university. These may include resiliency and problem-solving strategies (Cartrette and Melroe-Lehrman, 2012), greater self-efficacy, and higher-order learning (Kuh *et al.*, 2010; Robnett *et al.*, 2015), critical skills such as communication and teamwork (Carter *et al.*, 2016), a greater sense of identification with their field (Hurtado *et al.*, 2011), and a feeling of preparedness for graduate education (Hathaway *et al.*, 2002). Ultimately, these experiences can benefit low-SES students and students of color and reduce achievement gaps between these students and their traditionally privileged peers (Kinzie *et al.*, 2008).

In addition to supporting degree attainment generally, UR experiences can increase STEM degree attainment for marginalized students through economic, social, and cultural capital. The mentor–mentee relationships developed during UR can be essential to students' persistence and achievement in STEM (Haeger and Fresquez, 2016). Career guidance, expanding research skillsets, and broadening a social network are beneficial for students' belonging in STEM (Schwartz, 2012). Opportunities to learn about the research process could contribute to students' connection with STEM and to their perceptions of themselves as capable of pursuing and attaining their degree. It is evident that UR comes with multiple advantages (i.e., interest in graduate education, critical thinking skills, and navigating STEM environments) that boost STEM interest and persistence (Huggins *et al.*, 2019). While inequitable systemic structures are deeply embedded in the culture of universities, UR experiences could mitigate the personal and academic challenges that marginalized students are more at risk of encountering. In doing so, STEM graduation rates can increase and, as a result, broaden the representation of individuals with varying cultural and ethnic backgrounds in STEM occupations.

Despite the abundant research on the benefits of UR and its potential to have a compensatory effect, there has not been a systematic analysis of how UR might be used to increase degree completion and close equity gaps in STEM. Previous research and national calls for action (National Academies of Sciences, Engineering and Medicine, 2017) have highlighted the need for more direct measures of academic success and have encouraged employing methodologically different approaches to demonstrating the impact of UR. Haeger and Fresquez's (2016) quasiexperimental study on mentoring for inclusion demonstrated the positive relationship between mentored research and academic performance, but their study relied on data from a single institution. We aim to respond to those calls by replicating the post hoc, quasiexperimental study by Haeger and Fresquez (2016) on a direct measure of student success using a larger sample of data from six U.S. public universities.

To explore whether UR can increase a student's likelihood of graduating college and close equity gaps in STEM college completion, we will answer the following questions:

1. Are there disparities in who participates in UR?
2. Do students who participate in UR have an increased likelihood of graduating in 4 or 6 y when compared with like-peers and controlling for student background characteristics?
3. How does participation in an UR experience impact equity gaps and graduation rates for marginalized students in STEM (first-generation students, low-income students, and students of color)?

Exploring these questions will provide actionable guidance for faculty, staff, and policymakers working to advance equity in higher education and in broader U.S. society. This guidance is especially important, as college budgets constrict due to crises of lower enrolment or budget restrictions, and as institutions make crucial decisions about how to best support educational equity with limited resources.

METHODS

To systematically explore UR as a mechanism for educational equity, we have collected data from six public institutions in the

U.S. on UR participation, student characteristics, academic performance, and time-to-graduation. To demonstrate the potential variability of impact, we have selected institutions that vary in size, location, and population served. We utilize a post hoc, quasiexperimental design to reduce self-selection bias. Propensity score matching (PSM), a methodology recognized by the Department of Education's What Works Clearinghouse as means of simulating random assignment (Song and Herman, 2010), is used to create an appropriate comparison group of like-peers, allowing us to better understand the impact of participation in UR on the likelihood of degree attainment.

Our study aims to contribute to the understanding of the impact of UR by addressing common limitations in previous studies. Previous research on UR is often limited by at least one, if not several, of the following factors:

1. Measures: research using solely proxies for academic achievement rather than direct measures outcomes.
2. Generalizability: data are collected from a single institution/program.
3. Methodology: study design does not include an appropriate control/comparison group.

Research on UR has relied heavily on subjective measures including student self-reported learning and development. Though students' perceptions of their learning are unquestionably valuable, they are also heavily associated with satisfaction and are subjective measures of actual learning and development (Pike, 2011). Research is needed that utilizes more standardized measures and looks beyond self-reported data. The research that has used more direct measures has largely been focused on a single institution, making it unclear if the impact of that specific UR program is generalizable to other programs and campuses (Nagda *et al.*, 1998; Haeger and Fresquez, 2016).

A recent, multi-institution study looked at the relationships between high-impact practices, including UR, and campus level graduation rates (Johnson and Stage, 2018). Though this study included many institutions and a direct measure of success, it focused on the correlation between institutionalization of high-impact practices (ranging from: not available, offered, to required) and then the overall campus graduation rate. However, this approach does not account for variation in the type nor the extent of high-impact practice participation, particularly in the case of universities where such practices are offered but not required. Our study includes data from multiple campuses and focuses on the impact of participating in faculty-mentored research experiences on overall student success.

Finally, very few studies follow an experimental or quasiexperimental design that used randomly assigned control groups or statistical modelling to simulate random assignment with the exception of a few studies which only utilized single institution samples and were not replicated on other campuses (Nagda *et al.*, 1998; Pender *et al.*, 2010; Haeger and Fresquez, 2016). Without random assignment to UR and a control/comparison group or modelling to otherwise account for differences between students who self-select into UR and those that do not, it is not possible to tell if students are high achieving because they participated in UR or if high achieving students are simply more likely to self-select into participating in UR.

Our study aims to address these gaps in understanding by looking at a direct measure of student success, primarily the likelihood of graduating college, at six different institutions, and through a post hoc, quasiexperimental design. Though graduation does not account for the richness or the full breadth of the impact of UR, it is a critically import metric that has real-life career and financial implications for students.

To understand how undergraduate researchers (URs) differ from the general student population and how UR relates to timely graduation, as well as address limitations in previous research, we use data from multiple campuses, employ a post hoc quasiexperimental design, and control for students' background characteristics and prior academic performance. The study was submitted to the Institutional Review Board at California State University, Monterey Bay. The research was classified not as human subjects research because we were using deidentified institutional records with no means to identify individuals. Data from each institution was deidentified by that university's Institutional Research Office before inclusion in the multi-institution dataset.

Sample

We collected data from six public universities across the U.S. for a total sample of 120,308 students. We conducted analyses about the relationship between UR and graduation rates on a subsample of these data that includes only students whose entrance status was as a first-time, full-time student (excluding transfer students and part-time students), and who were enrolled in STEM majors by the end of their second year (n = 12,449; see Table 1 for subsample demographics). Of these students, 11% (n = 1426) were URs. The institutions all have established UR programs and were recruited through the Council on Undergraduate Research (CUR). Institutions were recruited through CUR at their annual business meeting and conference. Institutions were selected based on their ability to systematically identify students who participated mentored, out-of-class, UR experiences and to connect those data to institutional records. The participating institutions include University of Texas-Arlington, University of Wisconsin-Eau Claire, George Mason University, California State Polytechnic University-Pomona, California State University-Monterey Bay, and University of North Carolina at Greensboro. Data were collected from entering cohort years 2008–2011 in order to examine graduation rates at 4 and 6 y. The six universities in this sample differ in size, location, and student populations served including both Predominately White Institutions and Minority Serving Institutions, but they all share a commitment to UR and to engaging marginalized students in research opportunities.

We use institutional data from each university; therefore, the students' race/ethnicity and gender are what is reported in the institutional record and conform to Integrated Postsecondary Education Data System (IPEDS) standards. This means that there is very little missing data in the institutional records apart from parental education and prior academic performance that had a small amount of missing data. Cases with missing data were excluded from the matching and further analysis. In using institutional data, we acknowledge the limitation of simplifying racial and gender identity in this way because these data excludes any nonbinary gender identities or any changes in gender identity for transgender students.

TABLE 1. STEM sample descriptive statistics ($N = 12{,}449$) All student demographics are binary coded, 0 = no and 1 = yes so that the means can be converted to percentages for each group (i.e., a mean of 0.11 for participation in UR means that 11% of students participated in UR).

	N	Min.	Max	Mean	SD
White	12449	0	1	0.43	0.495
Black	12449	0	1	0.09	0.289
Asian/Pacific Islander	12449	0	1	0.18	0.384
Native American	12449	0	1	0.00	0.054
Latino/a/x	12449	0	1	0.17	0.378
Two or more races selected	12449	0	1	0.04	0.188
Gender Female	12446	0	1	0.44	0.496
First-generation in college	12210	0	1	0.28	0.450
Pell eligible ever	12449	0	1	0.39	0.487
Participation in UR	12449	0	1	0.11	0.318
Prior Academic Performance	12144	.0	4	3.18	0.60540

The designation of UR participation is based on records kept by the UR programs at each campus. For this study, UR is conceptualized as undergraduate participation in research, outside of class, and mentored by faculty, graduate students, or postdoctoral researchers. URs in the study received payment or academic credit for their participation; however, due to differences in data collection at each institution, we are not able to distinguish which students received funding and which received credit. Future studies should explore the differential impact of funded and credited or unfunded research. The scope of the current project does not include course-based research, but future studies should also examine that intervention.

We use Pell grant eligibility as a proxy for family income to account for the experience of lower-income students. The vast majority of Pell eligible students also received Pell grants, but because of differences in tracking between institutions, we only use Pell eligibility and not actual Pell recipient as our measure. Whether a student is a first-generation college student is used as a measure of socioeconomic status and parental education (Means and Pyne, 2017). This measure is based on the university's designation of the student as first-generation or not, and there is some variability with whether students whose parents attended college without graduating are counted as first-generation in college.

It is often assumed that higher performing students are more likely to participate in research both because of self-selection and faculty-mentor selection bias. For this reason, it was critical to include a measure of prior academic performance. We used the cumulative GPA after the first year of attendance as a measure of prior academic performance except for one campus where incoming GPA (i.e., high school GPA) was used instead.

A strength of this sample of institutions is that multiple campuses serve a diverse population of students. One campus is a Minority Serving Institution, and another serves a large proportion of low-income and first-generation students. The other campuses include programs to promote inclusion and diversity in UR resulting in a diverse population of researchers. Previous studies of UR have often been conducted at Predominantly White Institutions and have included larger proportions of traditionally privileged students (e.g., middle-to-upper-income, from college educated families, and White students; Haeger *et al.*, 2015). Having a more representatively diverse sample allows us to further test the impact of UR on all students and not just traditionally privileged students.

Post Hoc Quasiexperimental Design

A common critique of research on UR and student success is that there are inherent differences between students who self-select into research opportunities and those who do not. Since we cannot randomly assign students to participate in research experiences, we use a post hoc statistical method to create a group of like-peers – students that are similar to the population of URs (Dehejia and Wahba, 2002). We used PSM to calculate a student's likelihood of participating in UR and then find a comparison group of students who were just as likely to do engage in research but did not participate for various reasons (Rosenbaum and Rubin, 1985; Peikes *et al.*, 2008). To create this matched comparison group, we used SPSS software and replicated the matching procedures used in Haeger and Fresquez (2016).

To answer the first research question and determine how URs were similar or different from the general student population, we conducted t tests (see Table 2) between UR participants and the general population within each institution on: gender, race/ethnicity, parental education, Pell eligibility, transfer status, major, and prior academic performance.

We then conducted logistic models predicting probability of participating in UR at each institution. Any characteristic that was significantly related to UR participation was used in the PSM model for that institution (see Table 2 for *significant* variables that were used to calculate the propensity score). Prior academic performance was the only significant predictor of UR participation that was consistent across all six institutions. This variability illustrates the importance of calculating the propensity score nested within institution to create the strongest match for that specific student population.

A logistic model using these characteristics as covariates was used to create a propensity score for the probability of participating in UR at each institution. In this model, the dependent variable is the probability/odds of participating in UR for each student, given the chosen covariates (Rosenbaum and Rubin, 1985; Peikes *et al.*, 2008). Nearest neighbor matching was used to match students who conducted research with students who had not been in UR (Austin, 2014). This involved matching treatment group students (UR participants) with comparison groups students (students who did not participate in UR) by matching them based on propensity scores for optimal matching and only accepting differences between propensity scores

TABLE 2. Difference in URs and the general student population. The numbers represent the percentage of URs in each category minus the percentage of students who have not participated in research in each category, so that positive numbers mean that the category is overrepresented in UR. Negative numbers mean that those students are underrepresented in UR (* $p < 0.01$ and ** $p < 0.001$), items in bold are all statistically significant based on the mean comparison. [1]Difference in early career GPA instead of percentage. [2] University 1 provided incoming GPA (high school or transfer) instead of early GPA. The relationship was also significant and positive.

	Univ. 1	Univ. 2	Univ. 3	Univ.4	Univ. 5	Univ. 6
Enter Status, First-Time, First-Year	__8%**__	__8%*__	__28%**__	__–6%**__	–1%	__51%**__
STEM	__9%**__	__43%**__	4%	__32%**__	6%	__55%**__
White	0%	5%	9%**	1%	1%	–5%
Black	0%	0%	1%	1%	__–11%**__	__–6%*__
Asian	1%	__4%*__	__–4%*__	__–3%*__	5%	__7%*__
Native American	0%	1%	0%	0%	1%	0%
Latino/a/x	0%	–6%	2%	–1%	9%*	–3%
Multi-Racial	1%	__–3%*__	__1%*__	1%	0%	3%
Gender Female	2%	__–1%**__	__–8%**__	–1%	–5%	__–19%**__
Pell Eligible	1%	1%	0%	8%	0%	–1%
First-Generation	__3%*__	–2%	__–4%*__	8%**	0%	–1%
Prior Academic Performance[1]	__+**2__	__0.33**__	__0.11**__	__0.20**__	__0.20**__	__0.58**__

less than 0.000001. The *t* tests were conducted to test difference in the comparison and treatment groups on prior academic performance, parental education, Pell eligibility, race/ethnicity, gender, and entering status. Before matching, each university had a number of significant differences between UR student population and the general student population (see Table 2), but after matching, these differences were dramatically reduced and not statistically significant. Where there were differences between the treatment and comparison groups in different dimensions of identity, the treatment group tended to contain a relatively greater proportion of students who have a marginalized identity (i.e., more low-SES students, more first-gen students, more students of color) in the UR group. Therefore, if we assume that there is a general tendency for relatively more privileged students to be advantaged in college outcomes and for relatively less privileged students to be disadvantaged, we could expect that this pattern would err on the side of underestimating benefits of UR.

Once the matching was completed, data from all the universities was merged to create the final dataset for analysis. The matched sample includes a comparison group that is similar to the UR group, though the match is not exact. We were limited by the variables available from each campus and recognize that other factors may allow future researchers to create a better model for matching. Despite this fact, the matched sample still provides a similar group of like-peers for comparison. Most critically, the difference in prior academic performance is dramatically reduced from a GPA difference of 0.26 in the full sample to a difference of 0.01 in the matched sample.

Data Analysis

Independent sample *t* tests with Cohen's *d* calculations were conducted on the full, prematched dataset to answer the first research question of who participates in research and to test for differences in graduation rates. All further analyses were conducted on the matched and merged dataset. To answer the second research question, we conducted logistic regressions to estimate the impact of UR on the likelihood of graduating in STEM at 4 or 6 y, controlling for students' characteristics (e.g., race, gender, parental education, and Pell eligibility) and prior academic performance. All items in the regression were tested for inter-item collinearity with no substantive differences found, which include all significant correlations < 0.25. We used the Nagelkerke (1991) R^2 as a measure of model fit and report this in the regression analysis. R^2 was tested on student background characteristics and prior academic performance and compared with the model including UR to assess improvement in model fit (Menard, 2000). All regression models were improved with the inclusion of UR and the final, full models are presented in the findings section.

To answer the third research question, we also conducted descriptive analyses and logistic regressions on a subset of the matched dataset to assess the impact of participation in UR for students who have been marginalized in STEM. The subset of the sample included comparison and treatment sample students majoring in STEM fields and who belonged to one or more of these marginalized groups: students of color (Black, Latino/a/x, Native American, or multiracial), first-generation college students, and Pell eligible students [$n = 1254$]). These students continue to experience equity gaps as a consequence of ongoing inequities and marginalization inside and outside college. This analysis will further explore whether participating in UR can increase graduation rates and reduce equity gaps.

FINDINGS

To answer our first research question on whether there are disparities in who participates in UR, we conducted *t* tests on the full, prematched sample of STEM students (see Table 3).

Effect size (Cohen's *d*) calculated on the magnitude of the difference between URs and students who have not participated in research. To interpret effect size: ES < 0.20 trivial, 0.20 to 0.49 small, 0.50 to 0.79 medium, and 0.80 or greater is large (Fritz *et al.*, 2012).

We found significant gaps in participating in research in terms of race/ethnicity, parental education, income level, and prior academic performance (Table 3). The largest difference was in prior academic performance. URs had a higher average GPA when starting college (3.15 nonresearchers and 3.41 for researchers, effect size = 0.4, $t = 5.89$, $df = 1642$, $p < 0.001$). Differences in ethnicities was variable by campus (see Table 2),

TABLE 3. Research Question 1. Mean comparisons between UR and general student population (non-UR) in STEM ($n = 12{,}449$) on full, prematched sample of STEM students (**$p < 0.001$). All student demographics are binary coded, 0 = no and 1 = yes so that the means can be converted to percentages for each group (i.e., a mean of 0.43 [non-UR] and 0.46 [UR] for White means that 43% of the control group [non-UR] are identified as White compared with 46% of students in UR). All student demographic variables are presented as the percentage of representation in that category for the control and treatment groups.

	Average Representation		Average GPA		Cohen's *d*
	Non-UR	UR	Non-UR	UR	Effect Size
White	43%	46%			0.06
Black	10%**	4%**			0.19
Asian/Pacific Islander	18%	18%			0.00
Native American	0.3%	0.4%			0.00
Latino/a/x	17%**	20%**			0.08
Multi-racial	4%	4%			0.00
Gender (Female)	44%	42%			0.04
First Generation	27%**	34%**			0.15
Pell Eligible	38%***	42%**			0.08
Prior Academic Performance (GPA)			3.15**	3.41**	0.43

but the population of URs included more Latino/a/x students (effect size = 0.08, $t = -2.94$, $df = 1761$, $p < 0.001$) and fewer Black students (effect size = 0.19, $t = 9.01$, $df = 2297$, $p < 0.001$) across the institutions. These institutions also had higher proportions of first-generation and low-income students in UR than in the general student population.

Graduation Rates

To explore the second research question on the relationship between UR participation and likelihood of graduation, we conducted descriptive and regression analysis. In a descriptive analysis (*t* test) of the match sample, we see that STEM students who participated in UR graduated at dramatically and significantly higher rates than their matched peer group (see Figure 1). At 4 y, 39% of URs graduated compared with 23% of their like peers (4-y graduation: $t = 7.82$, $df = 2162$, $p < 0.001$). The increase in graduation rates was even more dramatic at 6 y with 95% of URs graduating compared with 56% of their peers (6-y graduation: $t = 25.39$, $df = 2162$, $p < 0.001$).

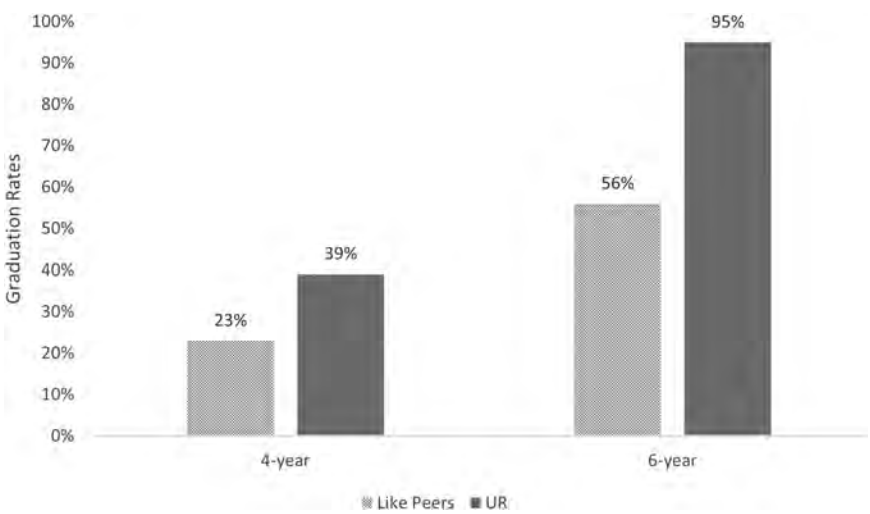

FIGURE 1. STEM graduation rate comparison between URs and like peers at 4- and 6-y ($n = 2727$).

To further test these patterns of higher graduation rates for UR participants, we conducted logistic regressions on the combined and matched dataset of STEM students for graduation in 4 and 6 y (Table 4).

The results of the regressions demonstrate that STEM students who have done UR are dramatically more likely to graduate within 4 and 6 y when compared with like peers and when controlling for background characteristics and prior academic performance. Students who participated in research were twice as likely to graduate in 4 y (Exp [β] = 1.964) and more than 10 times as likely to graduate in 6 y (Exp [β] = 11.295). The experimental betas for UR should not be directly compared with the experimental betas for other variables because using a matched sample may increase the variance in UR.

In addition to the dramatic impact of UR, a number of other factors were also significantly related to likelihood of graduation. Predictably, prior academic performance was the strongest, positive predictor of graduation. Though there were marginally significant differences by race/ethnicity in 4-year graduation rates, these were not significant at 6 y, and none of the interaction terms were significant.

Closing Equity Gaps in Graduation Rates

The previous analyses demonstrate that there are unequal patterns of participation in UR in terms of race/ethnicity and socioeconomic status (though these vary greatly by institution) and that participation in UR is related to increased likelihood of graduating college. To answer the third research question, the next set of analyses explores a subset of the data for students marginalized in STEM (students majoring in STEM fields who were Black, Latino/a/x, Native American, or multiracial, a first-generation college student, or were Pell eligible [$n = 1254$]). As seen in Figure 2, marginalized URs in STEM graduate at much higher rates than

TABLE 4. Logistic regressions on the likelihood of graduating in STEM (n = 2727; * p < 0.01 and ** p < 0.001). $_{1}$ Experimental Beta coefficients are presented as odds ratios so that values less than one are a reduction in likelihood of graduation and values over one are an increase in likelihood of gradation. The analyses included controls for institution attended, but the Exp (β) is not presented based on the data usage agreement with the universities. Reference group for race/ethnicity is White and the reference group for gender is male. [1]Interaction effect for student of color (Black, Native American, Latino/a/x, or multiracial).

	Grad 4 y R^2 0.08 Exp (β)	Grad 6 y R^2 0.358 Exp (β)
Black	0.300*	0.520
Asian/Pacific Islander	0.672*	1.280*
Native American	0.155	0.187
Latino/a/x	0.398*	0.689
Two or More Races	0.294*	0.420
Gender (Female)	1.104	0.854
Pell eligible ever	0.593	0.988
First-generation in College	1.033	1.042
Prior Academic Performance	1.591**	1.750**
Participation in UR	1.964**	11.295**
Student of color interaction effect$_1$	2.120	1.833
First-generation interaction effect	0.774	1.258
Pell eligibility interaction effect	1.467	1.642
Constant	0.486**	4.069**

other marginalized students in STEM. Forty-eight percent of marginalized URs graduated in 4 y compared with 35% of their like peers (4-y graduation: t = -10.66, df = 12545, p < 0.001). The increase in graduation rates was even more dramatic at 6 y (6-y graduation: t = -32.01, df = 12545, p < 0.001; Figure 2). Importantly, when we look at 6-y graduation rates, we see that 95% of URs marginalized in STEM graduate in 6 y and which is the same rate of graduation as the overall UR population. There is no equity gap in 6-y graduation rates, though a small gap (48 vs. 53%) persists in 6-y graduation rates (Figures 1 and 2).

When examining graduation rates for groups of marginalized students, we see dramatic equity gaps and striking differences between students who participate in UR and their like-peers (Figure 3). For students who have been historically excluded from research experiences based on race/ethnicity (Black, Latino/a/x, Native American, and multiracial students), we see a – 12% equity gap in 4-year graduation, but that equity gap is reduced to – 4% for URs (Figure 3).

We found a similar pattern in 4-y graduation equity gaps for Pell eligible and first-generation students with equity gaps cut in half for URs (Figure 3). At 6 y, the equity gap for students who have been excluded from research experiences based on race/ethnicity remains high (–10%) for non-UR students but is marginal for Pell eligible (–1%) and nonexistent (1%) for first-generation students. For URs at 6 y, the race/ethnicity equity gap is eliminated, and all marginalized groups had graduation rates slightly higher than average.

To further explore the role of UR in closing equity gaps, we conducted logistic regressions on the relationship between participation in UR and graduation (at 4 and 6 y) for students marginalized in STEM (e.g., students of color, first-generation students, and Pell eligible students; Table 5).

Similar to the regressions on all URs, in this analysis on students marginalized in STEM, we see that prior academic performance and participation in UR have strong and significant positive relationships to graduation. At 4 y, marginalized students are 1.69 times as likely to graduate if they participated in UR, and at 6 y, UR participants were a dramatic 14.36 times more likely to graduate.

DISCUSSION

This study explored how UR may be related to graduation rates and whether participation in research may reduce equity gaps in college graduation rates. To do this, we examined both the equity gaps that exist in who participates in research opportunities and how participation in UR is related to the likelihood of graduating when compared with like-peers. The results illustrate both promising signs of increasing participation in UR at some institutions and the importance of ensuring equitable participation because of the relationship between research participation and graduation rates.

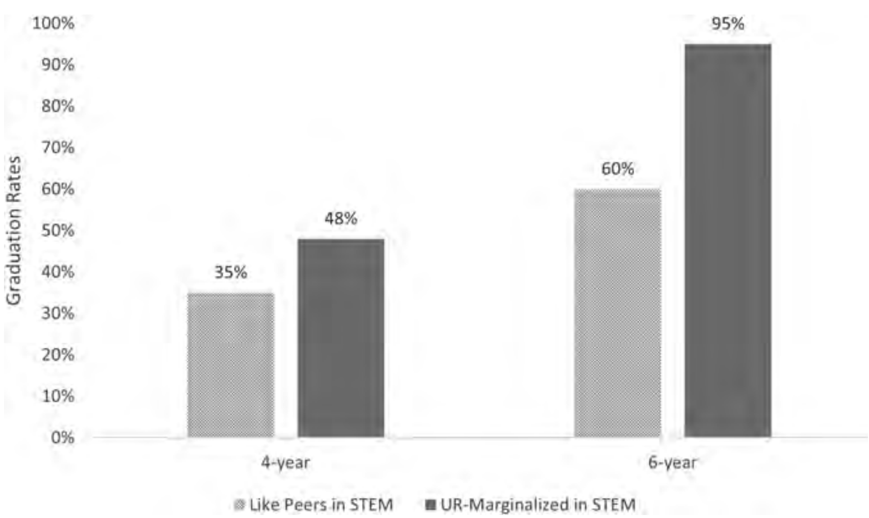

FIGURE 2. Graduation rate comparison between URs who have been marginalized in STEM and like-peers in STEM (n = 1254).

Who Participates in Research

In answering the first research question, we found that URs differed from the general student population in a number of ways. There were many more researchers in STEM majors and UR students had a higher first-year or transfer cumulative GPA on average than their peers. Current funding structures and faculty promotion standards in STEM fields are more conducive to engaging students in research in contrast to other disciplines that do not have access to the same level of federal grant funding to support UR (Ambos, 2020). Notably,

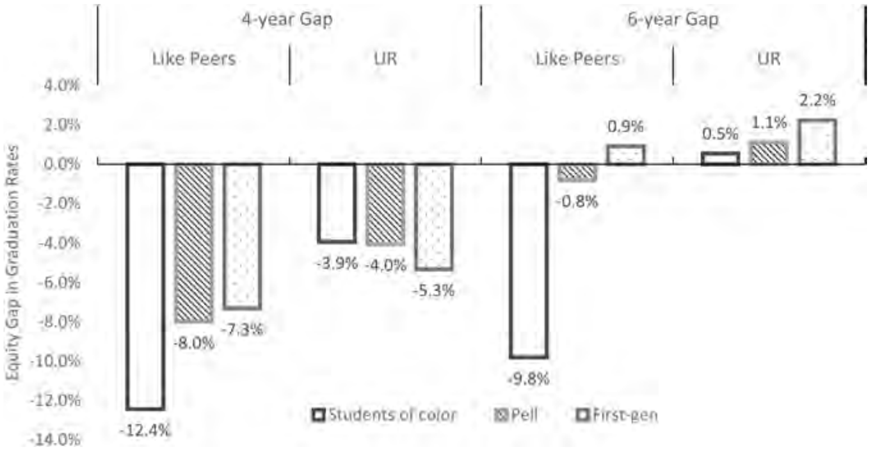

FIGURE 3. Equity gaps in STEM graduation rates for students of color, Pell eligible students, and first-generation students ($n = 2727$) for URs and like peers from the same marginalized group. Equity gaps are the differences between the average graduation rates and the marginalized group with the axis representing the average graduation rate.

though, the universities involved in this study all had highly diverse UR programs; indeed, at many of the institutions, the population of URs was actually significantly more diverse than the general student population in terms of parental education, student SES, and race/ethnicity. This is in stark contrast to previous research showing that marginalized students are less likely to have engaged in UR (National Academies of Sciences Engineering, and Medicine, 2017) even at minority serving institutions (Haeger *et al.*, 2015). Our finding suggests that programmatic efforts to increase diversity can have a dramatic impact in increasing participation in UR which is critical for improving persistence in STEM fields and diversifying the STEM workforce (Linn *et al.*, 2015)

TABLE 5. Logistic regressions on the likelihood of graduating in STEM fields conducted on a sub-sample of marginalized students ($n = 1254$; $*p < 0.01$, $** p < 0.001$, and $*** p < 0.0001$). Experimental Beta coefficients are presented as odds ratios so that values less than one are a reduction in likelihood of graduation and values over one are an increase in likelihood of gradation. The analyses included controls for institution attended, but the Exp (β) is not presented based on the data usage agreement with the universities.

	Grad 4 y R^2 0.202 Exp (β)	Grad 6 y R^2 0.174 Exp (β)
Black_1	0.716*	1.081
Asian/Pacific Islander	0.917	1.217*
Native American	0.835	0.778
Latino/a/x	0.738**	0.832*
Two or More Races	0.679*	0.745
Gender (Female)	1.242***	1.126**
Pell eligible ever	0.910*	1.043
First-generation in College	0.773***	0.804***
Prior Academic Performance	2.166***	1.759***
Participation in UR	1.686***	14.361***
Constant	0.128***	0.332***

UR and Student Success

This study has responded to calls for research on UR that can be generalized beyond a specific program or institution. We have responded to specific suggestions from the U.S. Department of Education's What Works Clearing House and the National Academies of Sciences, Engineering and Medicine (2017) to: 1) use more direct and objective measures of student success (graduation rates), 2) gather data from multiple institutions including MSIs, and 3) utilize analyses include an appropriate comparison group through experimental or quasi experimental designs. Through this study using more objective measures, data from six institutions, and with a post hoc, quasiexperimental design, we found a pattern of increased student success – operationalized as graduating from college – across multiple universities. This pattern suggests that our findings are generalizable beyond a specific UR program.

In using this multi-institution dataset to answer the second research question, we found that UR had a strong and significant relationship to a student's likelihood of graduating even when compared with like-peers and controlling for student background characteristics, institution attended, and prior academic performance. Previous research had suggested that high-impact practices including UR did not have a significant relationship to graduation rates (Johnson and Stage, 2018); however, that research did not differentiate between UR and other high-impact practices and focused on institutional policies (e.g., mandating vs. offering UR experiences). Such analyses may overlook the impact on individual students who participate, and our study suggests that is exactly where we see the dramatic difference that UR can make in students' academic success. Future research should examine what types of opportunities are offered by institutions which *require* all students to engage in research and how these opportunities may differ from individually mentored research experiences. Indeed, research has shown that factors such as the rigor and quality of the research experience along with access to financial and interpersonal support during research positively impact the outcomes students gain from the experience (Wilson *et al.*, 2018). Students who were in funded, rigorous research opportunities with access to faculty and peer support were more likely to go on to graduate school and produced more research products (e.g., posters and scholarly papers) than other students (Wilson *et al.*, 2018).

Though the present study demonstrates a positive relationship between participating in research and gradating college, we do not wish to suggest UR is a "magic bullet" which can solve most or all educational inequities. For instance, it is crucial to investigate the inequitable outcomes that may arise for marginalized populations even within the research context and even at racially/ethnically diverse institutions. UR experiences do not always result in positive outcomes given that mentoring interactions vary and have the potential to both hinder and foster students' research development (Aikens *et al.*, 2017).

Despite receiving one-on-one mentoring, first-generation STEM students are less likely to professionally advance (e.g., publish) compared with peers whose parents went to college, even when first-generation students are attending a more racially and socioeconomically diverse institution (Grineski *et al.*, 2017). Our findings suggest that involvement in UR is key for boosting graduation rates and broadening STEM participation. Yet, equity-focused researchers must also continue identifying factors that could impede students' professional development and potentially hinder their degree completion.

A common critique of associating UR with academic success is the assumption that already high achieving students are the students who engage in UR. We found that URs did have a higher average prior academic performance (3.04 first-year GPA for the nonresearchers and 3.33 GPA for UR). Despite this, we still found a dramatic, positive effect of participation when comparing to similarly higher achieving peers and when controlling for prior academic performance. This suggests that the higher rates of academic success for URs are not solely due to self-selection bias.

Beyond improving a student's likelihood of graduating at 4 and 6 y, the results from the third research question also demonstrate the potential of UR as a means of closing educational equity gaps. Prior research has demonstrated that equity gaps for student of color (Riegle-Crumb *et al.*, 2019) and lower-SES students (Bettencourt *et al.*, 2020) are larger in STEM fields. We found that equity gaps in 4-y graduation rates for students of color, low-income, and first-generation students were cut in half for URs. At 6-y, these gaps were completely closed for URs. Our results are consistent with prior research on UR and its relation to increasing STEM degree completion in marginalized student populations (Carpi *et al.*, 2017; Estrada *et al.*, 2016). While the quality of UR is not noted in the current study, our study replicated findings from studies at single institutions that participation in UR can increase academic success (Haeger and Fresquez, 2016) and reduce equity gaps (Bhattacharyya and Chan, 2021). Our results indicate that UR is a positive factor for STEM graduation rates across varying racial/ethnic minoritized student populations and for lower-SES students at different institutions. UR could be salient to developing a sense of belonging through multiple circumstances (e.g., connecting with STEM major through lab tasks, learning how to navigate the STEM environment, and building a STEM community through faculty and graduate student mentors; Olivares-Donoso and González, 2019), and thus contribute to diversifying the STEM field (Carpi *et al.*, 2017). As we seek ways to close educational equity gaps and increase graduation rates, UR can be a meaningful way to improve student success and increase equity.

ACKNOWLEDGMENTS

This research was partially funded through the Department of Education Hispanic Serving Institution (HSI) STEM and Articulation Programs Grants # P031C160221, Title: *Research-based Interventions to Increase STEM Degree Attainment*, and # P031C210011, Title: *Culturally Responsive Engagement, Articulation, and Research* (Project CREAR). The development of the multi-institution dataset was also supported by a collaborative project between California State University, Monterey Bay and the Council for UR.

Additional support was provided by the University of Arizona. Project CREAR (Culturally Responsive Engagement, Articulation and Research) which is funded through the U.S. Department of Education Hispanic Serving Institutions STEM and Articulation Program, Title III, Part F.

REFERENCES

Aikens, M. L., Sadselia, S., Watkins, K., Evans, M., Eby, L. T., & Dolan, E. L. (2016). A social capital perspective on the mentoring of undergraduate life science researchers: An empirical study of undergraduate-postgraduate-faculty triads. *CBE—Life Sciences Education*, *15*(2), 1–15. https://doi.org/10.1187/cbe.15-10-0208

Aikens, M. L., Robertson, M. M., Sadselia, S., Watkins, K., Evans, M., Runyon, C. R., ... & Herrera, J. (2017). Race and gender differences in undergraduate research mentoring structures and research outcomes. *CBE—Life Sciences Education*, *16*(2), 1–12. https://doi.org/10.1187/cbe.16-07-0211

Ambos, E. L. (2020). Undergraduate Research in the United States: Diversity, Growth, and Challenges. In Hensel, N. H., & Blessinger, P. (Eds.), *International Perspectives on Undergraduate Research* (pp. 19–38). Hampshire, UK: Palgrave Macmillan. https://doi.org/10.1007/978-3-030-53559-9_2

Artemeva, N. (2011). An engrained part of my career: The formation of a knowledge worker in the dual space of engineering knowledge and rhetorical process. *Writing in Knowledge Societies*, 321–350.

Austin, P. C. (2014). A comparison of 12 algorithms for matching on the propensity score. *Statistics in Medicine*, *33*, 1057–1069. https://doi.org/10.1002/sim.6004

Bettencourt, G. M., Manly, C. A., Kimball, E., & Wells, R. S. (2020). STEM degree completion and first-generation college students: A cumulative disadvantage approach to the outcomes gap. *The Review of Higher Education*, *43*(3), 753–779. https://doi.org/10.1353/rhe.2020.0006

Bhattacharyya, P., & Chan, C. W. M. (2021). Can Undergraduate Research Participation Reduce the Equity Gap? *Journal of the Scholarship of Teaching and Learning*, *21*(1), 287–300. https://doi.org/10.14434/josotl.v21i1.30462

Bourdieu, P. (1986). The forms of capital. In Richardson, J. (ed.), *Handbook of theory and research for the sociology of education* (pp. 241–258) Westport, CT: Greenwood.

Broton, K. M., & Goldrick-Rab, S. (2018). Going without: An exploration of food and housing insecurity among undergraduates. *Educational Researcher*, *47*(2), 121–133.

Cabrera, A. F., Stampen, J. O., & Hansen, W. L. (1990). Exploring the effects of ability to pay on persistence in college. *The Review of Higher Education*, *13*(3), 303–336.

Cabrera, A. F., & La Nasa, S. M. (2000). Understanding the college-choice process. *New Directions for Institutional Research*, *2000*(107), 5–22.

Carpi, A., Ronan, D. C., Falconer, H. M., & Lenis, N. H. (2017). Cultivating minority scientists: Undergraduate research increases self-efficacy and career ambitions for underrepresented students in STEM. *Journal of Research in Science Teaching*, *54*(3), 169–194. https://doi.org/10.1002/tea.21341

Carnevale, A. P., Rose, S., & Cheah, B. (2011). The college payoff: Education, occupation and lifetime earnings*. Center on Education and the Workforce*. Washington, DC: Georgetown University. Retrieved March 1, 2022, from www2.ed.gov/policy/highered/reg/hearulemaking/2011/collegepayoff.pdf

Carter, D. F., Ro, H. K., Alcott, B., & Lattuca, L. R. (2016). Undergraduate research experiences in promoting engineering students' communication, teamwork, and leadership skills. *Research in Higher Education*, *57*(3), 363–393.

Cartrette, D. P., & Melroe-Lehrman, B. (2012). Describing changes in undergraduate students' preconceptions of research activities. *Research in Science Education*, *42*(6), 1073–1100. https://doi.org/10.1007/s11165-011-9235-4

Chang, M. J., Eagan, M. K., Lin, M. H., & Hurtado, S. (2011). Considering the impact of racial stigmas and science identity: Persistence among biomedical and behavioral science aspirants. *Journal of Higher Education*, *82*(5), 564–596. https://doi.org/10.1080/00221546.2011.11777218

Cohen, E., Lotan, R., Scarloss, B., & Arellano, A. (1999). Complex instruction: Equity in cooperative learning classrooms. *Theory Into Practice*, *38*, 80–86. https://doi.org/10.1080/00405849909543836

Dehejia, R. H., & Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. *The Review of Economics and Statistics*, *84*(1), 151–161.

Eagan, Jr, M. K., Hurtado, S., Chang, M. J., Garcia, G. A., Herrera, F. A., & Garibay, J. C. (2013). Making a difference in science education: The impact of undergraduate research programs. *American Educational Research Journal*, *50*(4), 683–713.

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. https://doi.org/10.1187/cbe.16-01-0038

Fritz, C. O., Morris, P. E., & Richler, J. J. (2012). Effect size estimates: Current use, calculations, and interpretation. *Journal of Experimental Psychology: General*, *141*(1), 2–18. https://doi.org/10.1037/a0024338

Gasman, M., Nguyen, T.-H., & Conrad, C. F. (2015). Lives intertwined: A primer on the history and emergence of minority serving institutions. *Journal of Diversity in Higher Education*, *8*(2), 120–138. https://doi.org/10.1037/a0038386

Gilmore, J., Vieyra, M., Timmerman, B., & Maher, M. (2015). The relationship between undergraduate research participation and subsequent research performance of early career STEM graduate students. *Journal of Higher Education*, *86*(6), 834–863.

Goldrick-Rab, S. (2016). *Paying the price: College costs, financial aid, and the betrayal of the American dream*. Chicago, IL: University of Chicago Press.

Goldrick-Rab, S., Baker-Smith, C., Coca, V., Looker, E., & Williams, T. (2019). *College and university basic needs insecurity: A national #RealCollege survey report*. Philadelphia, PA: The Hope Center. https://hope4college.com/wp-content/uploads/2019/04/HOPE_realcollege_National_report_digital.pdf

Grineski, H., Daniels, H., Collins, T., Morales, D. X., Frederick, A., & Garica, M. (2017). The conundrum of social class: Disparities in publishing among STEM students in undergraduate research programs at a Hispanic majority institution. *Science Education*, *102*(2), 283–303. https://doi-org.ezproxy3.library.arizona.edu/10.1002/sce.21330

Hathaway, R. S., Nagda, B. R. A., & Gregerman, S. R. (2002). The relationship of undergraduate research participation to graduate and professional education pursuit: An empirical study. *Journal of College Student Development*, *43*(5), 614–631.

Haeger, H., BrckaLorenz, A., & Webber, K. (2015). Participation in undergraduate research at Minority Serving Institutions. *Perspectives on Undergraduate Research and Mentoring*, *4*(1), 1–22. http://blogs.elon.edu/purm/files/2015/11/Haeger-et-al-PURM-4.1-1.pdf

Haeger, H., & Fresquez, C. (2016). Mentoring for inclusion: The impact of mentoring on undergraduate researchers in the sciences. *CBE—Life Sciences Education*, *15*(3), ar36. https://doi.org/10.1187/cbe.16-01-0016

Hausmann, L. R. M., Feifei, Y., Ward Schofield, J., & Woods, R. L. (2009). Sense of belonging and persistence in White and African American first-year students. *Research in Higher Education*, *50*, 649–669.

Huggins, M. T., Molek, K. S., Cavnar, P., & Godfrey, L. (2019). Using scholars programs to enhance success of underrepresented students in Chemistry, Biomedical Sciences, and STEM. In: *Growing diverse stem communities: Methodology, impact, and evidence* (pp. 139–150). Washington, DC: American Chemical Society (ACS) Publications.

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.

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. https://psycnet.apa.org/doi/10.1353/csd.2012.0028

Johnson, S. R., & Stage, F. K. (2018). Academic engagement and student success: Do high-impact practices mean higher graduation rates? *The Journal of Higher Education*, *89*(5), 753–781. https://doi.org/10.1080/00221546.2018.1441107

Jones, M. T., Barlow, A. E., & Villarejo, M. (2010). Importance of undergraduate research for minority persistence and achievement in biology. *The Journal of Higher Education*, *81*(1), 82–115.

Kinzie, J., Gonyea, R., Shoup, R., & Kuh, G. D. (2008). Promoting persistence and success of underrepresented students: Lessons for teaching and learning. *New Directions for Teaching and Learning*, *2008*(115), 21–38. https://doi.org/10.1002/tl.323

Kiyama, J., & Rios-Aguilar, C. (2017). *Funds of knowledge in higher education: Honoring students' cultural experiences and resources as strengths*. New York, NY: Routledge.

Kuh, G. D., Kinzie, J., Schuh, J. H., & Whitt, E. J. (2010). *Student success in college: Creating conditions that matter,* 1st ed., San Francisco, CA: Jossey-Bass.

Linn, M. C., Palmer, E., Baranger, A., Gerard, E., & Stone, E. (2015). Undergraduate research experiences: Impacts and opportunities. *Science*, *347*(6222), 1261757. https://doi.org/10.1126/science.1261757

Manzoni, A., & Streib, J. (2019). The equalizing power of a college degree for first-generation college students: Disparities across institutions, majors, and achievement levels. *Research in Higher Education*, *60*(5), 577–605. https://doi.org/10.1007/s11162-018-9523-1

McClellan, G. S., Creager, K. L., & Savoca, M. (2018). *A good job: Campus employment as a high-impact practice*. Sterling, VA; Stylus Publishing, LLC.

Menard, S. (2000). Coefficients of Determination for Multiple Logistic Regression Analysis. *The American Statistician*, *54*(1), 17–24. https://doi.org/10.2307/2685605

Means, D. R., & Pyne, K. B. (2017). Finding my way: Perceptions of institutional support and belonging in low-income, first-generation, first-year college students. *Journal of College Student Development*, *58*(6), 907–924. https://doi.org/10.1353/csd.2017.0071

Nagda, B. A., Gregerman, S. R., Jonides, J., von Hippel, W., & Lerner, J. (1998). Undergraduate student faculty research partnerships affect student retention. *Review of Higher Education*, *22*(1), 55–72.

Nagelkerke, N. J. D. (1991). A note on a general definition of the coefficient of determination. *Biometrika*, *78*(3), 691–692. https://doi.org/10.1093/biomet/78.3.691

National Academies of Sciences, Engineering and Medicine. (2017). *Undergraduate research experiences for STEM students: Successes, challenges, and opportunities*. Washington, DC: The National Academies Press. https://doi.org/10.17226/24622

NCES (2019). *Indicator 23: Postsecondary graduation rates*. Retrieved March 8, 2022, from https://nces.ed.gov/programs/raceindicators/indicator_red.asp

Nunley, J. M., Pugh, A., Romero, N., & Seals, R. A. (2015). Racial discrimination in the labor market for recent college graduates: Evidence from a field experiment. *The B.E. Journal of Economic Analysis & Policy, De Gruyter*, *15*(3), 1093–1125.

Olivares-Donoso, R., & González, C. (2019). Undergraduate research or research-based courses: Which is most beneficial for science students? *Research in Science Education*, *49*(1), 91–107. https://doi.org/10.1007/s11165-017-9616-4

Peikes, D. N., Moreno, L., & Orzol, S. M. (2008). Propensity score matching. *The American Statistician*, *62*(3), 222–231. https://doi.org/10.1198/000313008x332016

Pender, M., Marcotte, D. E., Domingo, M. R. S., & Maton, K. I. (2010). The STEM pipeline: The role of summer research experiences in minority students' Ph.D. aspiration. *Education Policy Analysis Archives*, *18*(30), 1–36.

Perna, L. W. (2010). Understanding the working college student. *Academe*, *96*(4), 30–33.

Pike, G. (2011). Using college students' self-reported learning outcomes in scholarly research. *New Directions for Institutional Research*, *150*, 41–58.

Riegle-Crumb, C., King, B., & Irizarry, Y. (2019). Does STEM stand out? Examining racial/ethnic gaps in persistence across postsecondary fields. *Educational Researcher*, *48*(3), 133–144. https://doi.org/10.3102/0013189x19831006

Robnett, R. D., Chemers, M. M., & Zubriggen, E. L. (2015). Longitudinal associations among undergraduates' research experience, self-efficacy, and identity. *Journal of Research in Science Teaching*, *52*(6), 847–867. https://doi.org/10.1002/tea.21221

Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*, *39*(1), 33–38. https://doi.org/10.2307/2683903

Schwartz, J. (2012). Faculty as undergraduate research mentors for students of color: Taking into account the costs. *Science Education Policy*, *96*(3), 527–542. https://doi-org.ezproxy1.library.arizona.edu/10.1002/sce.21004

Smith, J. M., & Lucena, J. C. (2016). Invisible innovators: How low-income first-generation students use their funds of knowledge to belong in engineering. *Engineering Studies*, *8*(1), 1–26. https://doi.org/10.1080/19378629.2016.1155593

Stephens, N. M., Fryberg, S. A., Markus, H. R., & Johnson, C. S. (2012). Unseen disadvantage: How American universities' focus on independence undermines academic performance of first-generation college students. *Journal of Personality and Social Psychology*, *102*(6), 1178–1197. https://doi.org/10.1037/a0027143

Stephens, N. M., Hamedani, M. G., & Destin, M. (2014). Closing the social-class achievement gap: A difference-education intervention improves first-generation students' academic performance and all students' college transition. *Psychological Science*, *25*(4), 943–953. https://doi.org/10.1177/0956797613518349

Song, M., & Herman, R. (2010). Critical issues and common pitfalls in designing and conducting impact studies in education: Lessons learned from the What Works Clearinghouse (Phase I). *Educational Evaluation and Policy Analysis*, *32*(3), 351–371. https://doi.org/10.3102/0162373710373389

Thompson, J. J., & Jensen-Ryan, D. (2018). Becoming a "science person": Faculty recognition and the development of cultural capital in the context of undergraduate biology research. *CBE—Life Sciences Education*, *17*(4), ar62. https://doi.org/10.1187/cbe.17-11-0229

Thelin, J. R. (2011). *A history of American higher education*. Baltimore, MD: Johns Hopkins University Press.

Walpole, M. (2003). Socioeconomic status and college: How SES affects college experiences and outcomes. *The Review of Higher Education*, *27*(1), 45–73.

Walton, G. M., & Cohen, G. L. (2007). A question of belonging: Race, social fit, and achievement. *Journal of Personality and Social Psychology*, *92*(1), 82–96.

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.

Wilbur, T. G., & Roscigno, V. J. (2016). First-generation disadvantage and college enrollment/completion. *Socius*, *2*, 2378023116664351. https://doi.org/10.1177/2378023116664351

Wilkins-Yel, K. G., Arnold, A., & Bekki, J. (2022). "I can't push off my own mental health": Chilly STEM climates, mental health, and STEM persistence among Black, Latina, and White graduate women. *Sex Roles*, *86*, 208–232. https://doi.org/10.1007/s11199-021-01262-1

Wilson, A. E., Pollock, J., Billick, I., Domingo, C., Fernandez-Figueroa, E. G., Nagy, E., … & Summers, A. (2018). Assessing science training programs: Structured undergraduate research programs make a difference. *BioScience*, *68*(7), 529–534. https://doi.org/10.1093/biosci/biy052

Yee, A. (2016). The unwritten rules of engagement: Social class differences in undergraduates' academic strategies. *The Journal of Higher Education*, *87*(6), 831–858. https://doi.org/10.1080/00221546.2016.11780889

Yosso, T. J. (2005). Whose culture has capital? A critical race theory discussion of community cultural wealth. *Race Ethnicity and Education*, *8*(1), 69–91. https://doi.org/10.1080/1361332052000341006