AMERICAN SOCIETY FOR MICROBIOLOGY | JMBE Journal of Microbiology & Biology Education

Research Article

Influence of CUREs on STEM retention depends on demographic identities

Lisa Bradshaw,[1] Julianne Vernon,[2] Thomas Schmidt,[1,3] Timothy James,[1] Jianzhi Zhang,[1] Hilary Archbold,[4] Kenneth Cadigan,[4] John P. Wolfe,[5] Deborah Goldberg[1]

AUTHOR AFFILIATIONS See affiliation list on p. 13.

ABSTRACT Research has shown that undergraduate research experiences can have substantive effects on retaining students in science, technology, engineering and mathematics (STEM). However, it is impossible to provide individual research experiences for every undergraduate student, especially at large universities. Course-based undergraduate research experiences (CUREs) have become a common approach to introduce large numbers of students to research. We investigated whether a one-semester CURE that replaced a traditional introductory biology laboratory course could increase retention in STEM as well as intention to remain in STEM, if the results differed according to demography, and investigated the possible motivational factors that might mediate such an effect. Under the umbrella of the Authentic Research Connection (ARC) program, we used institutional and survey data from nine semesters and compared ARC participants to non-participants, who applied to ARC but either were not randomly selected or were selected but chose not to enroll in an ARC section. We found that ARC had significant effects on demographic groups historically less likely to be retained in STEM: ARC participation resulted in narrowing the gaps in graduation rates in STEM (first vs continuing-generation college students) and in intention to major in STEM [females vs males, Persons Excluded because of Ethnicity or Race (PEERs) vs non-PEERs]. These disproportionate boosts in intending STEM majors among ARC students coincide with their reporting a greater sense of student cohesiveness, retaining more interest in biology, and commenting more frequently that the course provided a useful/valuable learning experience. Our results indicate that CUREs can be a valuable tool for eliminating inequities in STEM participation, and we make several recommendations for further research.

KEYWORDS undergraduate research, CURE, STEM retention, introductory biology laboratory

Multiple studies show that a large percentage of undergraduate students who initially intend to major in a science, technology, engineering, or mathematics (STEM) discipline do not finish with a STEM degree (1–3). Although the most recent compilation of statistics indicates that switching from STEM to non-STEM majors declined from 1991 to 2011, many students who enter university intending a STEM major still do not finish with a STEM degree (4). Additionally, the attrition rates among demographic groups with historically lower representation in STEM continue to be higher than those of overrepresented groups: more women than men leave STEM during their undergraduate career (4). Black, Latinx, and Native American students had lower 4-, 5-, and 6-year undergraduate completion rates in STEM compared to Asian American and white students (5). Bettencourt et al. (6) found first-generation students were less likely to complete a STEM (or a non-STEM) degree compared to students who had at least one parent with a college degree.

Editor Samantha T. Parks, Georgia State University, Atlanta, Georgia, USA

Address correspondence to Deborah Goldberg, degold@umich.edu.

The authors declare no conflict of interest.

Received 25 November 2022
Accepted 7 September 2023
Published 25 October 2023

In their landmark study "Talking About Leaving," Seymour and Hewitt (7) documented 24 issues that students identified as contributing to their decision to switch from STEM majors, and these issues were also of concern to students who persisted in STEM majors. Two decades later, students reported the same issues, but the average number of concerns reported by switching students increased from 4.2 to 12 (8). A little over one third of those issues reference negative aspects of students' STEM classroom experiences, while some others include financial issues, loss of confidence, loss of subject interest, competitive STEM culture, difficult transition to college, and inadequate high school preparation (8). Furthermore, these factors accumulate and often interact with demographic traits which make a more complex and nuanced explanation of how/why students leave or persist in STEM. For instance, Canning et al., (9) documented that first-generation students' perceptions of their STEM classroom as highly competitive correlated with more frequent intentions to drop the STEM course compared to continuing-generation students with the same competitive assessment.

One of the solutions to increase STEM retention with demonstrated success is early participation in research (10, 11), but typically, low faculty-to-student ratios limit the number of individually mentored research projects available to a small fraction of students. The financial cost per student in many instances also prohibits offering such research experiences more widely. Course-based undergraduate research experiences (CUREs) provide a more scalable way to offer undergraduate students research experience. CUREs as defined by Auchincloss et al. (12) have become widespread among universities as a means of introducing STEM students to the process of "real science" (13–16). Furthermore, due to their scalability, CUREs can be more inclusive to help increase the numbers of historically underrepresented groups in STEM (17, 18).

The most direct way to assess CUREs' impacts on STEM persistence is to compare graduation rates in STEM for students who had a CURE experience with those who did not. The few studies taking this approach are positive (19, 20; Supplement 1), as are results from shorter term assessments that ask students about their intended major pre- and post-CURE (17, 21; Supplement 1). However, limited data have addressed the impact of demographic identity with CURE experiences and, therefore how CUREs might reduce gaps in STEM participation..

In addition to trends in STEM majors, it is important to understand how CUREs mediate students' decisions to stay in STEM. At the time we began our study in 2015, several proposed frameworks for evaluating the short- and long-term effectiveness of CUREs had been published based upon previous studies and the wider literature on learning theory assessment and psychology instruments (12, 22, 23). Broadly categorized, the models coalesce on measuring students' self-efficacy, attitude/motivation, and science identity/sense of belonging to a science community as key outcomes. A number of studies have shown increases in one or two of these motivational factors with participation in CUREs (16, 18, 24–30; Supplement 1), but few have considered their relative importance.

In this study, we tested the hypotheses that participation in a semester-long CURE would lead to higher graduation rates with a STEM major, increased intention of a STEM major, as well as changes in student motivation and perception of the classroom environment, with particularly strong impacts on groups historically excluded from STEM. To determine student motivation and perception, we drew from published survey instruments to construct a tool for assessing effort belief (31) and laboratory confidence (24) as measures of self-efficacy; maintained subject interest (31, 32) and perceived usefulness and importance (33) as measures of attitude; and student cohesiveness as a proxy for science identity (34, 35). We also included intellectual accessibility (33) as a proxy for both self-efficacy and attitude. We ask the following:

1. Did students who participated in a semester-long CURE in an introductory biology laboratory course graduate with STEM degrees at a higher rate than students in a traditional laboratory course and does this depend on demography?
2. Did students in the CURE biology laboratory remain intent on or decide on a STEM major by the end of the semester at a higher rate than students in a traditional laboratory course and does this depend on demography?
3. Did students' self-reported attitude, self-efficacy, and perceptions of the classroom environment differ between a CURE and the traditional laboratory course and among demographic groups? How do any differences relate to the graduation and intended major data?

METHODS

A comprehensive description of the methods, including statistical approaches, can be found in Supplement 2.

The course

Biology 173 is a two-credit laboratory course required for all biology majors at the University of Michigan Ann Arbor campus (UM), as a prerequisite for many upper-level laboratory courses, and for non-majors planning to apply to medical school. Approximately 1,500 students enroll in it every year, most of them during their second year. We recruited biology faculty to design and teach sections of Bio 173 as a CURE, with a semester-long research project based on their own laboratory's research topics, which we refer to as streams. Thus, the CUREs in our study were novel and designed to contribute to the faculty members' research programs. (For a detailed description of course format and research streams, see Supplement 2.)

Student participation

We advertised the research streams under the umbrella of the Authentic Research Connection program (ARC), directing students to the ARC website and application link. We randomly selected from the pool of first- and second-year student applicants to fill 75% of available seats and then randomly selected from the entire applicant pool to fill the remaining seats, with some exceptions as noted in Supplement 2. In every semester except winter 2015, more students applied than the streams could accommodate. Students who applied but were not selected or who were offered a spot but did not enroll in the research streams for any reason we designate as "non-participants" and differentiate in the analyses from non-applicants, i.e., students who did not apply to ARC streams. The non-participants serve as the comparison group to ARC participants to account for any self-selection bias or other differences between students who applied to ARC versus those who were not interested or aware of the option.

Surveys

Beginning in Fall 2015, we surveyed all students enrolled in Bio 173 during the first 2 weeks (pre-semester) and final 2 weeks (post-semester) of the course. Requests to participate and informed consent for participation in the surveys were conducted in accordance with institution norms and under approval by the University of Michigan IRB (study ID HUM00094780). In addition to asking for demographic data and intended major, we asked students about five motivational factors and one classroom environment factor, as described in more detail in (30) and Supplement 3.

Graduation in STEM

We searched university records to find all students who had applied to participate in ARC for fall 2015 and onward and had graduated as of May 2020 and recorded their majors and degree awarded. As the non-applicant group for graduated students, we used students who took Bio 173 in fall 2015 but had not applied to ARC. We used the US Department of Education Classification of Instructional Programs codes STEM Designated Degree Program (https://www.ice.gov/sites/default/files/documents/stem-list.pdf, accessed 3/30/2021) list to classify students as obtaining a STEM or non-STEM degree, with exceptions as noted in Supplement 2.

Data analysis

We categorized students into streams as denominated in Table 1. In some cases, the ARC streams were lumped for analysis because of small sample size. For demographic variables, we classified gender as male or female (too few students used a nonbinary descriptor for statistical analysis); ethnicity as Persons Excluded because of their Ethnicity or Race [PEER, defined as persons who identify as Black or African American, Latinx, or Hispanic and peoples indigenous to the US and its territories (36)] or non-PEER; and parental college history as first generation or continuing generation. All statistical analyses except the two-way and three-way tests of independence were conducted in the R environment using R studio (RStudio Team, 2020). For all analyses, sample sizes were too small to examine the interactions among the demographic groupings such as gender × ethnicity, even though those intersections are increasingly recognized as critical (4, 6).

To compare pre-semester intended majors with either post-semester intended major (STEM, other, undecided) or major at graduation (STEM, other, no degree), we first used two-way χ^2 tests of independence to look for overall differences at each stage and among levels of each demographic grouping. We then used three-way log-linear analyses (G-tests) to examine how intended pre-semester and post-semester or actual majors depended on the interaction of demography and research stream.

For motivational and classroom environment factors, we calculated gain scores as a student's mean post-semester score minus their mean pre-score. We ran multiple regression models with gain scores as the dependent variable and pre-scores, stream, demographic variables, the interactions of pre-score with stream and with each demographic variable, and the interactions of stream with each demographic variable as

TABLE 1 Details of Biology 173 ARC (CURE) and regular sections during the study period[f]

Stream of Biology 173 lab course	No. of semesters analyzed	Term and yr of semesters offered	No. of 1st- and 2nd-year student responses in analysis	No. of students graduated included in degree analysis
CURE, Fly genetics	1	fall 18	24	0
CURE, Yeast evolution	2	win 18, win 19, win 20[a]	69	12
CURE, Human microbiome	9	win 15[b], fall 15, win 16, fall 16, win 17, fall 17, win 18, fall 18, win 19, fall 19[c], win 20	509	393
Regular, Non-participant	9	fall 15, win 16, fall 16, win 17, fall 17, win 18, fall 18, win 19, fall 19, win 20	322	178
Regular, Non-applicant	9	fall 15, win 16, fall 16, win 17, fall 17, win 18, fall 18, win 19, fall 19, win 20	324[d]	455[e]

[a] The emergence of the COVID-19 pandemic during the winter 2020 semester caused all courses to move to an online format halfway through the semester, so data from this semester were excluded from analysis.

[b] The survey instrument was not developed and validated until the summer of 2015; thus, no survey data from winter 2015 exist.

[c] Student cohesiveness questions were omitted from the post semester survey in fall 2019, which decreased the number of total responses for that factor only.

[d] Subset of responses randomly selected for analysis from the larger set (3,319) of all non-applicant survey respondents.

[e] All graduated non-applicant students in the analysis were enrolled in Biology 173 in Fall 2015 semester.

[f] Non-participant describes students who applied to participate in ARC but were not offered a spot or did not enroll in an ARC stream. Non-applicant describes students enrolled in the regular course sections who did not apply to participate in ARC. Note that sample sizes for analyses are often slightly smaller than the numbers indicated here because of missing data from surveys on intended major or demography.

the predictors. We then compared hierarchical models to determine the most parsimonious, best-fit model.

Grades are thought to correlate positively with students' perceptions of course quality and motivational factors. Therefore, ideally, we would have incorporated individual grades as a potential factor influencing survey results, along with stream and demography. However, we did not track individual grades during the study so were unable to explicitly account for grades. Instead, we obtained de-identified grade records with demographic data for all sections of Bio 173 from all semesters during the study period; analyses and results are described in Supplement 4.

Finally, we compiled the written comments from all post-semester surveys along with the stream and semester when they were made, using word frequency analysis in QDA Miner (Provalis Research, 2020) to refine an initial category list. We used the frequency analysis function in QDA Miner to determine the counts and percent frequencies of comments in each category for each stream and then performed two-way χ^2 tests to compare frequencies of comments among the streams.

RESULTS

Throughout our analyses, we frequently found differences between non-participants and non-applicants, which affirms that differences existed between those students who applied to ARC and those who did not. Results for the non-applicants are shown in tables and figures, but we limit our discussion to the more appropriate comparison of non-participants and ARC participants, who all asked to participate in a CURE.

Graduation rates in STEM

Average graduation rate in STEM, regardless of participation in ARC, among students who enrolled in Bio 173 was high, although we also found previously documented demographic discrepancies in retention (Table 2). Overall, participation in ARC did not affect graduation with a STEM degree (Table 2). However, the lack of overall effects of ARC on graduation in STEM masks some interesting interactions of stream with demography on STEM majors. Most notably, for parental college history, more continuing-generation students than first-generation students obtained STEM degrees for all streams, but the gap was smaller for ARC participants than for non-participants (the significant three-way interaction; Fig. 1C). For ethnicity, more non-PEERs than PEERs

TABLE 2 Results of χ^2 tests on distribution of intended majors and degree types by demographic groups and by streams for students who had graduated as of May 2020[a]

		Intended major					Degree obtained				
Group		**% STEM**	**% Other**	**% Undecided**	**χ^2**		**% STEM**	**% Other**	**% No degree**	**χ^2**	
Gender											
n = 708	Female	67	23	10	**10.415**	**	68	30	2	**12.270**	**
n = 329	Male	76	14	10			77	20	2		
Ethnicity											
n = 103	PEER	66	24	10	1.235		62	34	4	5.645	0.06
n = 934	Non-PEER	70	29	10			72	26	2		
Family											
n = 143	First generation	70	24	6	4.358		61	32	7	**26.597**	***
n = 894	Continuing generation	70	20	10			73	26	1		
Stream											
n = 405	ARC (CURE)	71	21	8	**16.108**	**	72	26	2	6.761	
n = 178	Non-participant	78	16	6			77	22	1		
n = 455	Non-applicant	66	21	13			69	26	2		

[a] ARC is consolidated as a single stream due to small sample sizes in Yeast evolution and Fly genetics (12 and 0, respectively). Non-participant indicates students who applied to participate in ARC but enrolled in regular Bio 173, and Non-applicant refers to students who enrolled in regular Bio 173 but did not apply to ARC. Overall graduation rate in STEM among students who enrolled in Bio 173 was >70% and just 2% of all students did not complete their degree at UM. Demographic composition of graduated students was as follows: female, 68%; male, 32%; PEER, 10%; non-PEER, 90%; first generation, 14%; continuing generation, 86%. Asterisks indicate significance at 0.05 (*), 0.01 (**), and 0.001 (***) levels. For marginally significant results, *P*-values are given.

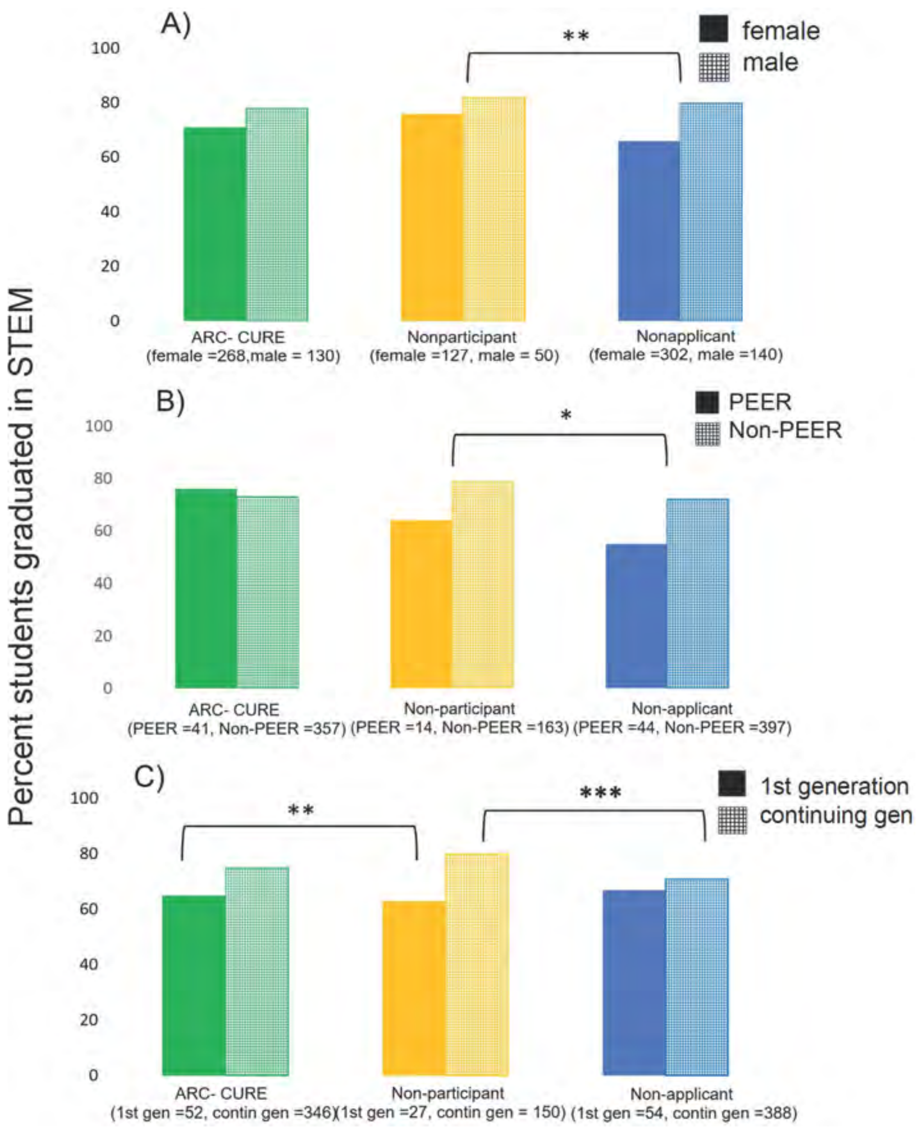

FIG 1 Percentage of students earning STEM degrees. Statistical results indicate the significance of the three-way interaction between stream (ARC vs non-participant or non-participant vs non-applicant) and demographic group (A, gender; B, ethnicity; C, parental college) and degree (STEM vs non-STEM). Number of students in each group shown in parentheses. "No degree" students were excluded, and ARC shows all CURE streams combined due to small number of graduates from the Yeast evolution and Fly genetics streams. Significance levels: *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$.

seemed to obtain STEM degrees for non-participants, but the gap disappeared for ARC participants (Fig. 1B). However, despite the apparently large effect size, the interaction of stream (ARC vs non-participants) with ethnicity (PEER vs non-PEER) was not significant, perhaps due to the small sample size ($n = 14$) of PEER non-participants. Gender did not influence effects of ARC on graduation with a STEM degree, although more males than females obtained STEM degrees for all streams (Fig. 1A).

STEM persistence before and after students' introductory biology laboratory semester

Regardless of participation in ARC, this larger data set showed similar results to those for graduation for comparisons by gender and ethnicity, although results differed somewhat for parental college history (Table 3). However, participation in ARC often interacted

TABLE 3 Results of χ^2 tests on distribution of intended majors pre- and post-semester by demographic groups and streams for first and second year students[a]

Group		Pre-semester intended major					Post-semester intended major				
		% STEM	% Other	% Undecided	χ^2		% STEM	% Other	% Undecided	χ^2	
Gender											
n = 841	Female	74	14	12	**7.227**	*	74	16	10	**11.287**	**
n = 359	Male	79	9	12			81	8	11		
Ethnicity											
n = 101	PEER	71	12	17	2.548		76	13	11	0.197	
n = 1099	Non-PEER	76	12	12			76	13	11		
Family											
n = 158	First generation	82	13	5	**6.912**	*	81	12	7	1.796	
n = 1042	Continuing generation	75	12	13			75	14	11		
Stream											
n = 461	Human microbiome	75	16	9	**26.808**	***	78	13	9	**20.81**	**
n = 69	Yeast evolution	85	6	9			87	7	6		
n = 24	Fly genetics	92	4	4			83	4	13		
n = 322	Non-participant	80	9	11			79	12	9		
n = 324	Non-applicant	69	13	18			68	17	15		

[a] Demographic composition of students was as follows: female, 70%; male, 30%; PEER, 8%; non-PEER, 92%; first generation, 13%; continuing generation, 87%. Asterisks indicate significance at 0.05 (*), 0.01 (**), and 0.001 (***) levels.

significantly with demographic categories. For gender, at the beginning of the semester, the gap between STEM intended males and females was marginally larger for ARC students than non-participants (Fig. 2A), but by the end of the semester, the gender gap for ARC participants had narrowed while widening for non-participants, resulting in no significant difference (Fig. 2B). The ethnicity gap between STEM intended non-PEERs and PEERs was smaller for ARC students at the beginning of the semester (Fig. 2C) but narrowed even more, from 9 to 5 percentage points, over the semester, and remained significantly different from non-participants' gap which went from 12 to 11 percentage points (Fig. 2D). For parental college status, no differences between ARC students and non-participants were found either pre- or post-semester.

Student motivation and classroom environment responses

To understand the relationship between students' Bio 173 experience and STEM persistence, we analyzed changes between their pre- and post-semester responses on survey questions relating to the five motivational factors and one classroom environment factor (Table S1). A positive gain score (post-semester score − pre-semester score) means an increase in the factor.

Analyses of variance (ANOVAs) indicated that pre-semester scores on the motivational factors often differed among demographic groupings and streams (Table S2). Females reported a greater interest in biology and rated its importance and usefulness higher, while males rated biology as more intellectually accessible and had greater laboratory confidence in pre-semester surveys. PEER and non-PEER students had similar pre-semester scores for all factors. First-generation students began the semester less confident in their laboratory skills than continuing-generation students but did not differ in any other factor. We also found some differences in pre-semester scores among streams (Table S2).

We only included significant ($P < 0.05$) predictors in Table 4. Best-fit models explained 23%–40% of the variance in the motivational and classroom environment factors (Table 4). Pre-semester scores of all factors were highly significant predictors in the models; students with higher initial factor scores increased less over the semester in that factor as indicated by the negative value of the coefficients.

Surprisingly, overall, gain scores were negative for subject interest, effort belief, and intellectual accessibility (Table S1); gain scores were positive for laboratory confidence and student cohesiveness (Table S1). The regression models showed that variation in gain scores were at least partially explained by stream and demography. Focusing

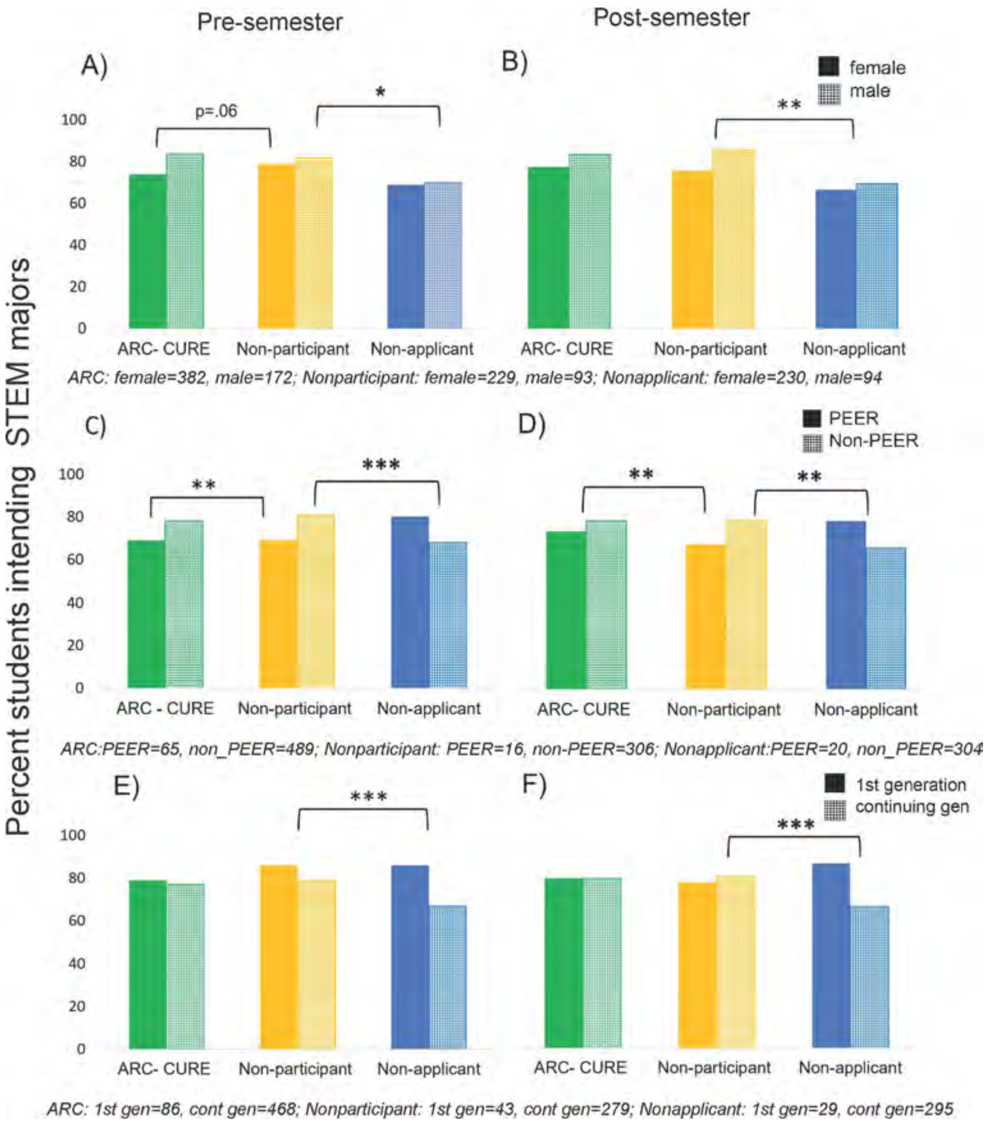

FIG 2 Percentage of students intending STEM majors by stream pre-semester and post-semester broken down by gender, ethnicity, and parental college history. Statistical results indicate the significance of the three-way interaction between stream (ARC vs non-participant or non-participant vs non-applicant) and demographic group (A and B, gender; C and D, ethnicity; E and F, parental college) and major (STEM vs non-STEM). Number of students in each group shown in parentheses. All ARC streams were combined due to small numbers of observations in sub-groups of Yeast evolution and Fly genetics. Significance levels: *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, *P*-values written above bars for marginal results.

on differences between non-participants and the three ARC streams, we found that stream was a significant predictor of students' gain scores for all factors except effort belief and interacted with gender for intellectual accessibility and with ethnicity for laboratory confidence (Table 4). Students in the Human microbiome stream had greater increases in subject interest scores (less negative) compared to non-participants (base level for model). On the other hand, Fly genetics students had greater decreases in subject interest, intellectual accessibility, and importance and usefulness. In laboratory confidence, although Yeast evolution as a single factor showed a negative impact on gain scores, its positive interaction with ethnicity offsets the effect. Human microbiome and Yeast evolution students both had smaller gaps in gains between PEER and non-PEER students in laboratory confidence relative to non-participants. Finally, all ARC streams showed greater gains in student cohesiveness compared to non-participants.

TABLE 4 Regression coefficients for best-fit models predicting the change in factor scores pre- to post-semester[a]

	Factor response variable					
	Subject interest	**Effort belief**	**Intellectual accessibility**	**Importance and usefulness**	**Lab confidence**	**Student cohesiveness**
Predictor variable ↓						
Pre-semester score	0.509 ***	0.493 ***	0.639 ***	0.545 ***	0.814 ***	0.639 ***
Stream: HM (CURE)	0.326 ***					0.173 ***
Stream: YE (CURE)					0.594 *	0.284 ***
Stream: FG (CURE)	0.491 **		0.312 *	0.238 *		0.234 $P = 0.06$
Stream: Non-applicant	0.173 **			0.126 **		
Gender			0.149 ***			
Ethnicity	0.141 *	0.100 *			1.060 ***	
Parental college						
Pre-score × streams (all)						
Pre-score × gender						
Pre-score × ethnicity					0.196 **	
Pre-score × parental college						
Stream HM CURE × gender						
Stream YE CURE × gender			0.332 *			
Stream FG CURE × gender						
Stream Non-applicant × gender	0.217 *					
Stream HM CURE × ethnicity					0.383 *	
Stream YE CURE × ethnicity					0.641 *	
Stream FG CURE × ethnicity						
Stream Non-applicant × ethnicity						
Streams (all) × parental college						
Adjusted R^2	**0.23**	**0.24**	**0.26**	**0.29**	**0.40**	**0.29**
df/error df	11/1219	6/1224	12/1218	5/1225	11/1219	8/1142

[a] Interactions among the demographic variables and all three-way interactions were not tested. The base levels of the model were as follows: non-participant, female, PEER, and first generation. Positive coefficients indicate that the designated levels of the categorical predictor variables had a greater increase in score relative to the base. Key to streams and number of students in ( ): HM, Human microbiome (509); YE, Yeast evolution (69); FG, Fly genetics (24); Non-participant (314); Non-applicant (315). Significance levels: *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$. Non-significant values not reported.

Gender only significantly affected intellectual accessibility gain scores, with females showing smaller gains than males, and this disparity was greater for Yeast evolution (Table 4). Ethnicity impacted three of the factors: PEER student scores decreased more than non-PEER students in subject interest and effort belief but increased more in laboratory confidence. Also, a significant interaction between ethnicity and pre-scores in laboratory confidence shows PEER students with lower laboratory confidence at the beginning of the semester gained more than non-PEER students, but those with higher confidence at the beginning showed similar gains. Parental college history by itself had no effect on gain scores but did show a significant interaction with pre-scores in intellectual accessibility, where first-generation students with higher pre-scores gained less than continuing-generation students with high pre-scores, while gains for low pre-score students were similar between first and continuing generations.

Post-semester survey comments

χ^2 Analysis across the four streams showed significant differences in every one of the nine categories of comments, so we then performed pairwise comparisons among streams. Non-applicants and non-participants did not differ significantly in their frequencies of comments in any category, indicating that non-participants did not form a more negative opinion of the regular course despite having applied to participate in an ARC section (Fig. 3). Compared to non-participants, Human microbiome participants made fewer complaints about grading, fewer negative comments about the instructor, the course overall, and about the work being tedious (Fig. 3). Conversely, they made

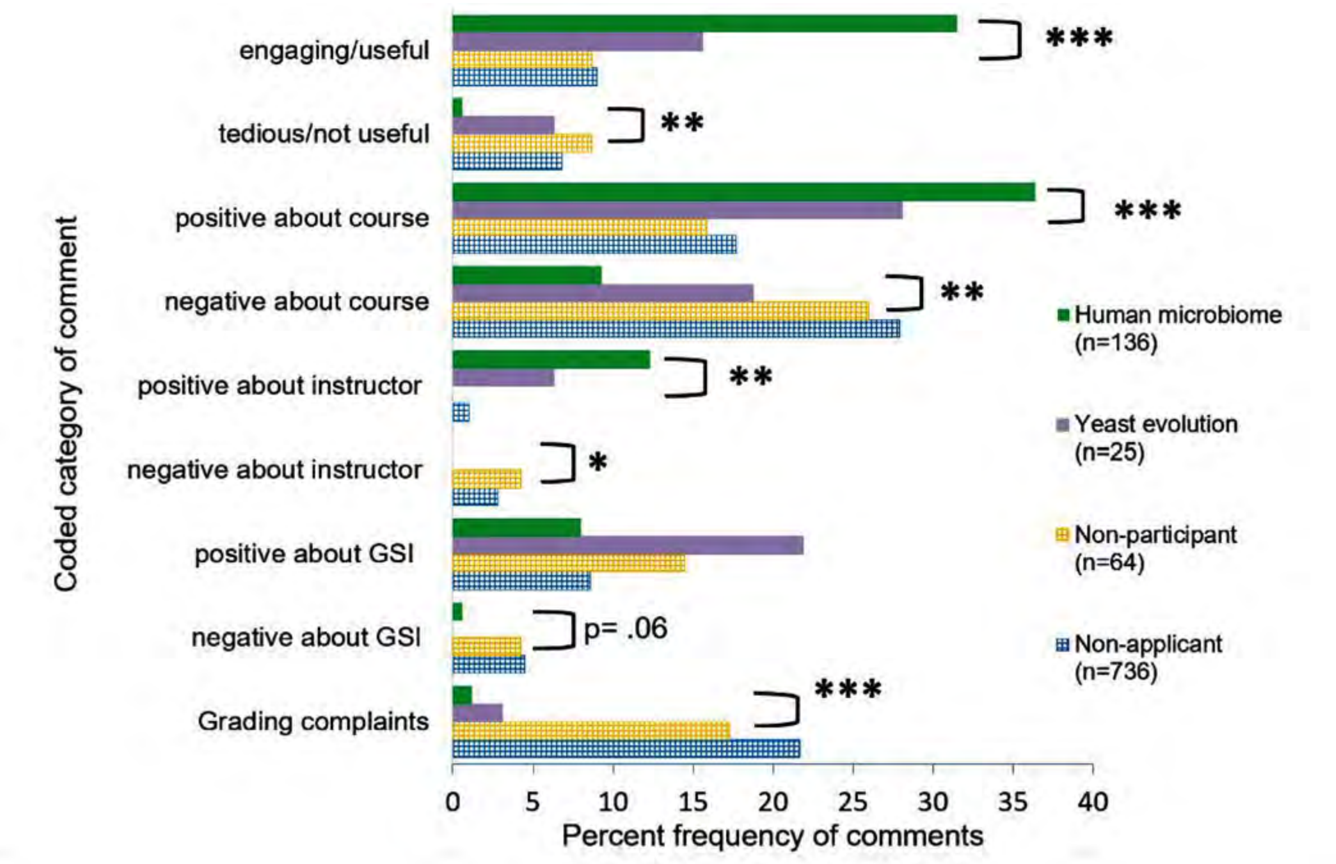

FIG 3 Relative frequencies of categorized survey comments by stream. All student comments in every stream were included. Significance levels indicate differences between Human microbiome CURE and non-participants (*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$). Yeast evolution CURE had more positive instructor comments ($P = 0.02$), marginally more positive course comments ($P = 0.056$), and marginally fewer grading complaints ($P = 0.07$) than non-participants. Compared to the Human microbiome CURE, the Yeast evolution CURE had more positive comments about the GSIs and more comments about the laboratory being tedious. There were no significant differences between non-participant and non-applicant comments.

more positive comments about the instructor and course overall and more frequently described the course as engaging/useful (Fig. 3). With only 25 student comments from the Yeast evolution stream, comparisons are less conclusive, but they made more positive comments about the instructor and marginally tended to make more positive comments about the course overall and less complaints about grading relative to non-participants. Interestingly, there were no negative comments about the instructor in either of the ARC streams and only a single negative comment about the graduate student instructor (GSI) in those streams. Too few comments were available for the Fly genetics ARC stream for statistical analysis.

DISCUSSION

Our study showed much higher degree completion rates and somewhat higher STEM retention rates at the University of Michigan Ann Arbor compared to national rates reported in references (1) and (5), although a lower STEM degree rate than participants in a three-semester-long CURE program at the University of Texas (19). We also note that a large majority of the students in our study were not in the first semester of their first year when enrolling in the introductory biology laboratory, so our sample population mostly did not capture those students who switch out of STEM in their first semester. Nevertheless, we found the same gaps in STEM retention between historically excluded demographic groups (women, PEERs, and first-generation students) and their overrepresented counterparts as other studies (4–6).

When averaging over all students, our data indicate no or relatively small effects of participating in a one-semester CURE (ARC) on graduation with a STEM degree or intentions to major in STEM. However, this overall result masks important differences among demographic groups in response to ARC. For groups that are typically less likely to be retained in STEM, participating in ARC resulted in narrowing the gaps either in graduation rates in STEM (first- vs continuing-generation college students) or in intention to major in STEM (females vs males, PEERs vs non-PEERs). These disproportionate boosts in intending STEM majors among ARC students in those demographic groups coincide with ARC students overall reporting a greater sense of student cohesiveness, making more positive comments about the instructor and course, and in the Human microbiome stream, retaining more interest in biology and commenting that the course provided a useful/valuable learning experience.

Below, we explore these results and their implications in more detail and then discuss several important potential limitations of this study.

Interactions of demography and effect of CUREs

Our results suggest that the ARC CURE experience was most beneficial for students who have been historically excluded from STEM fields, specifically women and PEER students. We see that the gender gap existed pre-semester, consistent with the fact that women rated biology as less intellectually accessible than males did pre-semester, and gender was a significant predictor of gain score (males out gain females) even after taking pre-score into account. However, for ARC students, the gender gap on STEM intention narrowed, while widening among non-participants pre- to post-semester, indicating women in ARC were more drawn to STEM by the experience. As a possible explanation, all ARC streams were significant predictors of higher gain scores in student cohesiveness which is a measure of student cooperation and support. Seymour and Hewitt (7) noted several studies showing women prefer cooperative learning strategies, and they found, as did Hunter (8), that women more frequently mention the competitive culture in STEM classes as discouraging and a reason for switching majors. Furthermore, women more frequently mention instructor's personal attention to them as important to their perception of good teaching (37), and PEER women in particular feel more comfortable approaching teaching assistants over professors for help (38). The fact that ARC streams had undergraduate instructional aides and a course liaison (usually a postdoc or technician from the instructor's laboratory group) in addition to GSIs for laboratory sections may have resulted in women in ARC feeling more supported, which positively impacted their inclinations toward STEM. For example, a Human microbiome female participant wrote "I feel that it is important to mention that I didn't have the best experience with my GSI, as he was sometimes very condescending and not helpful, which inhibited a collaborative and open learning environment. Had the other GSI and professors not been so helpful and accommodating, I think that this particular experience would have stayed with me and affected my choices of whether or not to enroll … in an additional [biology] course." Loss of interest in the subject matter was a top 3 factor in students' decision to leave STEM (8), and while this did not differ between men and women, students in the Human microbiome ARC maintained more subject interest in contrast to non-participants. For female ARC students, this maintained interest may have had an additive effect.

Although the difference was small, a higher percentage of PEERs than non-PEERs in ARC finished with STEM degrees, and the ethnicity gap in percent STEM-intended closed from pre- to post-semester for ARC students, indicating movement towards STEM for ARC PEERs. This appears contradictory to our survey results showing that PEER students overall experienced greater decreases in subject interest and effort belief over the semester compared to non-PEERs, factors which parallel some of the top reasons students give for switching out of STEM highlighted earlier in this paper. However, we also found that PEER students in ARC streams made greater relative gains in laboratory confidence (a measure of self-efficacy) compared to non-participant PEERs. The greater student cohesiveness reported by ARC students maps to the long-term outcome of identifying as a scientist, which Estrada et al. (39, 40) found more highly correlated with scientific integration and STEM persistence, respectively, in a study of PEER students. A similar study of students of color majoring in STEM at a predominantly white institution documented peer group support as a main theme in those students' explanations of their persistence in STEM (41). Additionally, Human microbiome students more frequently described the course as an engaging/useful learning experience, and students in both ARC streams more often made general positive comments about the course, which maps to making connections and greater student engagement, components associated with positive STEM course experiences. Thus, it appears these positive aspects of ARC offset the negative effects experienced by PEER students in subject interest and effort belief, leading them into STEM majors.

ARC's impact on first-generation students is less clear. While the generation gap in STEM degree completion was smaller for ARC, no patterns emerged in the pre- and post-semester intended majors, and parental college did not predict gain scores for any of the survey factors. Previous studies have shown that first-generation students' lower STEM retention relative to continuing-generation students is driven by different factors such as lower average family income, math scores, and high school preparation (6, 42), resulting in what Bettencourt et al. (6) term a "cumulative disadvantage." Notably though, Stephens et al. (43) documented that, irrespective of intended major, first-generation students had more communal, interdependent motives for attending college, while continuing-generation students had more independent motives, and the independent culture of US universities presents another disadvantage for first-generation students. Participating in the Human microbiome CURE which ties into a larger research program with implications for human health may have reinforced these community-centered motives and thus increased the appeal of a STEM major. Consider the sentiment and use of plural pronouns in this comment from a first-generation student in the Human microbiome stream: "We're, in a way, on the front line of this type of gut microbiome research - and it's really exciting. The professors and GSIs are all really passionate about the coursework/research, and that makes it easy to be excited about class." The greater sense of student cohesiveness reported by ARC participants, which aligns with communal values, may also be relevant for increasing STEM persistence among first-generation students.

Potential limitations and directions for future research

Where possible, we analyzed the research streams separately and found they sometimes differed in their comparisons to non-participants. We suggest these differences may be attributed to the fact that we only had data from a single, pilot semester of Fly genetics and the first two semesters of Yeast evolution, compared to nine semesters which did not include the first run for Human microbiome. In their debut semesters, Yeast evolution and Fly genetics experienced typical difficulties with first-run laboratory experiments which may have impacted student affect and comments about the course. Our inconsistent results (except for student cohesiveness) from the streams which only ran for one or two semesters further underscore the importance of conducting multi-year studies before drawing conclusions about a program's effectiveness.

One of the strengths of our study is that we randomly selected students to participate in the ARC sections from among those who applied, which allows us stronger inference about the influence of participation in a CURE than if all students self-selected into the CURE. However, not all students selected to participate in ARC ended up enrolling in an ARC section and we do not have data on the causes of this lack of participation. Thus, we do not have a fully randomized experiment. In addition, five semesters after the project began when we began to see clear benefits of participation, we started to admit all PEER students who applied into an ARC, increasing the social benefit of our program, but further reducing our ability to infer causality for this demographic group.

Performing more quantitative validity tests, such as factor analysis [as explained in reference (44)], could reduce uncertainty that the motivation and classroom experience factors in our survey map directly to the broader constructs of self-efficacy, attitude, and science identity discussed above and, therefore strengthen interpretation from our regression analyses.

Another potential limitation is that we were unable to explicitly account for the impact of grade differentials between CURE and regular streams (see Supplement 4 for further discussion).

Finally, an important caveat is that uncontrolled-for differences in instruction may have impacted our results. Most switchers in "Talking About Leaving Revisited" listed poor STEM teaching and course design as contributing to their decision to leave STEM majors (8). In our study, students in the regular Bio 173 course more frequently made negative comments about the course, the instructor, and the GSIs and complained of unfair grading, high workload, and perceived disorganization. With an enrollment of over 700 students per semester, Bio 173 requires at least 15 different GSIs to lead all the laboratory sections, and they have very different levels of interest and experience in teaching, as well as varying expertise in the subdisciplines covered. In contrast, all instructors in the ARC streams, from the undergraduate aides to the faculty instructor, had experience and interest in the course research topic which they could transfer to the students. We have no way to isolate the effects of individual instructors' teaching style and rapport with students but note that the trends we found are compiled from nine semesters of instruction, five different faculty course instructors, and dozens of GSIs and other classroom support personnel. In addition to quality of instruction, as noted above, ARC sections have more instructional staff for each section. While difficult and expensive to implement, future studies could control for experience, knowledge, number, and type of instructional staff. Alternatively, qualitative methods such as interviews with students would help validate results. Post exit interviews with those students who switch intended majors during the semester could provide especially salient information about how their course experience influenced their degree plans.

Conclusions

We found that a one-semester CURE led to increased STEM retention among demographic groups who have historically been excluded, and this impact appears most strongly related to perceived greater student cohesiveness in the classroom, with higher maintained subject interest and finding the course engaging/useful also contributing. While previous students have shown that CUREs can increase retention in STEM, they have not tested for differential effects among demographic groups and thus that CUREs can help close the gap in representation in STEM. We suggest future studies include more questions measuring science identity/sense of belonging to explore its short-term direct and longer-term mediated impact. We also would advise researchers to focus on long-running CUREs for more reliable results. Greater control of quality of instruction between CUREs and regular sections would also be useful to strengthen inference about the mechanisms driving impacts of CUREs on STEM retention.

ACKNOWLEDGMENTS

We thank the University of Michigan students who participated in this study, as well as the teaching assistants, graduate student instructors, and liaisons who supported the ARC program.

Support for this work was provided by a grant from the Howard Hughes Medical Institute (52008119).

AUTHOR AFFILIATIONS

[1]Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, Michigan, USA

[2]Dean's Office, School of Engineering, Vanderbilt University, Nashville, Tennessee, USA
[3]Division of Infectious Diseases, Department of Internal Medicine, University of Michigan, Ann Arbor, Michigan, USA
[4]Department of Molecular, Cellular and Developmental Biology, University of Michigan, Ann Arbor, Michigan, USA
[5]Chemistry Department, University of Michigan, Ann Arbor, Michigan, USA

AUTHOR ORCIDs

Thomas Schmidt http://orcid.org/0000-0002-8209-6055
Timothy James http://orcid.org/0000-0002-1123-5986
Deborah Goldberg http://orcid.org/0000-0002-7950-0432

AUTHOR CONTRIBUTIONS

Lisa Bradshaw, Conceptualization, Formal analysis, Investigation, Methodology, Project administration, Visualization, Writing – original draft | Julianne Vernon, Conceptualization, Investigation, Methodology, Project administration, Writing – review and editing | Thomas Schmidt, Conceptualization, Writing – review and editing | Timothy James, Conceptualization, Writing – review and editing | Jianzhi Zhang, Conceptualization, Writing – review and editing | Hilary Archbold, Conceptualization, Writing – review and editing | Kenneth Cadigan, Conceptualization, Writing – review and editing | John P. Wolfe, Conceptualization, Funding acquisition, Writing – review and editing | Deborah Goldberg, Conceptualization, Funding acquisition, Methodology, Supervision, Writing – review and editing

ADDITIONAL FILES

The following material is available online.

Supplemental Material

Supplement 1 (jmbe00225-22-s0001.docx). Detailed literature review.
Supplement 2 (jmbe00225-22-s0002.docx). Detailed methods.
Supplement 3 (jmbe00225-22-s0003.docx). Survey questions.
Supplement 4 (jmbe00225-22-s0004.docx). Impact of grades.
Supplemental Table S1 (jmbe00225-22-s0005.pdf). Gain scores.
Supplemental Table S2 (jmbe00225-22-s0006.pdf). Prescores.

REFERENCES

1. Chen X. 2013. STEM attrition: College students' paths into and out of STEM fields (NCES 2014-001). Washingon, DC US Department of Education
2. Hayes RQ, Whalen SK, Cannon B. 2009. CSRDE STEM retention report (center for institutional data exchange and analysis). University of Oklahoma, Norman, OK.
3. Toven-Lindsey B, Levis-Fitzgerald M, Barber PH, Hasson T. 2015. Increasing persistence in undergraduate science majors: a model for institutional support of underrepresented students. CBE Life Sci Educ 14:ar12. https://doi.org/10.1187/cbe.14-05-0082
4. Weston TJ. 2019. Patterns of switching and relocation, p 55–85. In Seymour E, AB Hunter (ed), Talking about leaving Revisited: Persistence, relocation, and loss in undergraduate STEM education. Springer, Cham, Switzerland.
5. Eagan K, Hurtado S, Figueroa T, Hughes B. 2014. Examining STEM pathways among students who begin college at four-year institutions. National Academy of Sciences, Washington, DC. https://sites.nationalacademies.org/cs/groups/dbassesite/documents/webpage/dbasse_088834.pdf.
6. Bettencourt GM, Manly CA, Kimball E, Wells RS. 2020. STEM degree completion and first-generation college students: a cumulative disadvantage approach to the outcomes gap. Rev High Ed 43:753–779. https://doi.org/10.1353/rhe.2020.0006
7. Seymour E, Hewitt NM. 1997. Talking about leaving: why undergraduates leave the sciences. Westview Press, Boulder, CO.
8. Hunter A-B. 2019. Why undergraduates leave STEM majors: Changes over the last two decades, p 87–114. In Seymour E, AB Hunter (ed), Talking about leaving Revisited: Persistence, relocation, and loss in undergraduate STEM education. Springer, Cham, Switzerland.
9. Canning EA, LaCosse J, Kroeper KM, Murphy MC. 2020. Feeling like an imposter: the effect of perceived classroom competition on the daily psychological experiences of first-generation college students. Soc Psychol Personal Sci 11:647–657. https://doi.org/10.1177/1948550619882032
10. Lopatto D. 2004. Survey of undergraduate research experiences (SURE): first findings. Cell Biol Educ 3:270–277. https://doi.org/10.1187/cbe.04-07-0045
11. Russell SH, Hancock MP, McCullough J. 2007. Benefits of undergraduate research experiences. Science 316:548–549. https://doi.org/10.1126/science.1140384
12. Auchincloss LC, Laursen SL, Branchaw JL, Eagan K, Graham M, Hanauer DI, Lawrie G, McLinn CM, Pelaez N, Rowland S, Towns M, Trautmann NM,

Varma-Nelson P, Weston TJ, Dolan EL. 2014. Assessment of course-based undergraduate research experiences: a meeting report. CBE Life Sci Educ 13:29–40. https://doi.org/10.1187/cbe.14-01-0004

13. Bakshi A, Patrick LE, Wischusen EW. 2016. A framework for implementing course-based undergraduate research experiences (cures) in freshman biology labs. Am Biol Teach 78:448–455. https://doi.org/10.1525/abt.2016.78.6.448
14. Harrison M, Dunbar D, Ratmansky L, Boyd K, Lopatto D. 2011. Classroom-based science research at the introductory level: changes in career choices and attitude. CBE Life Sci Educ 10:279–286. https://doi.org/10.1187/cbe.10-12-0151
15. Jordan TC, Burnett SH, Carson S, Caruso SM, Clase K, DeJong RJ, Dennehy JJ, Denver DR, Dunbar D, Elgin SCR, Findley AM, Gissendanner CR, Golebiewska UP, Guild N, Hartzog GA, Grillo WH, Hollowell GP, Hughes LE, Johnson A, King RA, Lewis LO, Li W, Rosenzweig F, Rubin MR, Saha MS, Sandoz J, Shaffer CD, Taylor B, Temple L, Vazquez E, Ware VC, Barker LP, Bradley KW, Jacobs-Sera D, Pope WH, Russell DA, Cresawn SG, Lopatto D, Bailey CP, Hatfull GF. 2014. A broadly implementable research course in phage discovery and genomics for first-year undergraduate students. mBio 5:e01051-13. https://doi.org/10.1128/mBio.01051-13
16. Shapiro C, Moberg-Parker J, Toma S, Ayon C, Zimmerman H, Roth-Johnson EA, Hancock SP, Levis-Fitzgerald M, Sanders ER. 2015. Comparing the impact of course-based and apprentice-based research experiences in a life science laboratory curriculum. J Microbiol Biol Educ 16:186–197. https://doi.org/10.1128/jmbe.v16i2.1045
17. Bangera G, Brownell SE. 2014. Course-based undergraduate research experiences can make scientific research more inclusive. CBE Life Sci Educ 13:602–606. https://doi.org/10.1187/cbe.14-06-0099
18. Graham MJ, Frederick J, Byars-Winston A, Hunter A-B, Handelsman J. 2013. Increasing persistence of college students in STEM. Science 341:1455–1456. https://doi.org/10.1126/science.1240487
19. Rodenbusch SE, Hernandez PR, Simmons SL, Dolan EL. 2016. Early engagement in course-based research increases graduation rates and completion of science, engineering, and mathematics degrees. CBE Life Sci Educ 15:ar20. https://doi.org/10.1187/cbe.16-03-0117
20. Indorf JL, Weremijewicz J, Janos DP, Gaines MS. 2019. Adding authenticity to inquiry in a first-year, research-based, biology laboratory course. CBE Life Sci Educ 18:ar38. https://doi.org/10.1187/cbe.18-07-0126
21. Shuster MI, Curtiss J, Wright TF, Champion C, Sharifi M, Bosland J. 2019. Implementing and evaluating a course-based undergraduate research experience (CURE) at a Hispanic-serving institution. Interdiscip J Probl Based Learn 13:1–13. https://doi.org/10.7771/1541-5015.1806
22. Brownell SE, Kloser MJ. 2015. Toward a conceptual framework for measuring the effectiveness of course-based undergraduate research experiences in undergraduate biology. Stud High Educ 40:525–544. https://doi.org/10.1080/03075079.2015.1004234
23. Corwin LA, Graham MJ, Dolan EL. 2015. Modeling course-based undergraduate research experiences: an agenda for future research and evaluation. CBE Life Sci Educ 14:es1. https://doi.org/10.1187/cbe.14-10-0167
24. Brownell SE, Kloser MJ, Fukami T, Shavelson R. 2012. Undergraduate biology lab courses: comparing the impact of traditionally based “cookbook” and authentic research-based courses on student lab experiences. J Coll Sci Teach 41(4):36–45.
25. Kloser MJ, Brownell SE, Shavelson RJ, Fukami T. 2013. Effects of a research-based ecology lab course: a study of nonvolunteer achievement, self-confidence, and perception of lab course purpose. J Coll Sci Teach 42:72–81.
26. Esparza D, Wagler AE, Olimpo JT, Hatfull GF. 2020. Characterization of instructor and student behaviors in CURE and non-CURE learning environments: impacts on student motivation, science identity development, and perceptions of the laboratory experience. CBE Life Sci Educ 19:ar10. https://doi.org/10.1187/cbe.19-04-0082
27. Newell MJ, Ulrich PN. 2022. Gains in scientific identity, scientific self-efficacy, and career intent distinguish upper-level cures from traditional experiences in the classroom. J Microbiol Biol Educ 23:e00051-22. https://doi.org/10.1128/jmbe.00051-22
28. Lopatto D, Alvarez C, Barnard D, Chandrasekaran C, Chung H-M, Du C, Eckdahl T, Goodman AL, Hauser C, Jones CJ, Kopp OR, Kuleck GA, McNeil G, Morris R, Myka JL, Nagengast A, Overvoorde PJ, Poet JL, Reed K, Regisford G, Revie D, Rosenwald A, Saville K, Shaw M, Skuse GR, Smith C, Smith M, Spratt M, Stamm J, Thompson JS, Wilson BA, Witkowski C, Youngblom J, Leung W, Shaffer CD, Buhler J, Mardis E, Elgin SCR. 2008. Genomics education partnership. Science 322:684–685. https://doi.org/10.1126/science.1165351
29. May NW, McNamara SM, Wang S, Kolesar KR, Vernon J, Wolfe JP, Goldberg D, Pratt KA. 2018. Polar plunge: semester-long snow chemistry research in the general chemistry laboratory. J Chem Educ 95:543–552. https://doi.org/10.1021/acs.jchemed.7b00823
30. Vernon J, Goldberg DE, Wolfe JP. 2016. “Engaging students in authentic research in introductory chemistry and biology Laboratories” 2016 ASEE Annual Conference & Exposition; New Orleans, Louisiana: https://doi.org/10.18260/p.26973
31. Ferrell B, Barbera J. 2015. Analysis of students' self-efficacy, interest, and effort beliefs in general chemistry. Chem Educ Res Pract 16:318–337. https://doi.org/10.1039/C4RP00152D
32. Harackiewicz JM, Durik AM, Barron KE, Linnenbrink-Garcia L, Tauer JM. 2008. The role of achievement goals in the development of interest: reciprocal relations between achievement goals, interest, and performance. J Educ Psychol 100:105–122. https://doi.org/10.1037/0022-0663.100.1.105
33. Bauer CF. 2008. Attitude toward chemistry: a semantic differential instrument for assessing curriculum impacts. J Chem Educ 85:1440. https://doi.org/10.1021/ed085p1440
34. Fraser BJ, Giddings GJ, McRobbie CJ. 1992. Assessing the climate of science laboratory classes. Key Centre for School Science and Mathematics, Perth, Australia. http://eric.ed.gov/?id=ED369657.
35. Dorman JP. 2003. Cross-national validation of the what is happening in this class? (WIHIC) questionnaire using confirmatory factor analysis. Learn Environ Res 6:231–245. https://doi.org/10.1023/A:1027355123577
36. Asai DJ. 2020. Race matters. Cell 181:754–757. https://doi.org/10.1016/j.cell.2020.03.044
37. Harper RP, Weston TJ, Seymour E. 2019a. Students' perceptions of good STEM teaching, p 245–276. In Seymour E, AB Hunter (ed), Talking about leaving Revisited: Persistence, relocation, and loss in undergraduate STEM education. Springer, Cham, Switzerland.
38. Harper RP, Weston TJ, Seymour E. 2019b. Student responses to problematic STEM teaching methods, p 149–196. In Seymour E, AB Hunter (ed), Talking about leaving Revisited: Persistence, relocation, and loss in undergraduate STEM education. Springer, Cham, Switzerland.
39. Estrada M, Woodcock A, Hernandez PR, Schultz PW. 2011. Toward a model of social influence that explains minority student integration into the scientific community. J Educ Psychol 103:206–222. https://doi.org/10.1037/a0020743
40. Estrada M, Hernandez PR, Schultz PW. 2018. A longitudinal study of how quality mentorship and research experience integrate underrepresented minorities into STEM careers. CBE Life Sci Educ 17:ar9. https://doi.org/10.1187/cbe.17-04-0066
41. PalmerRT, Maramba DC, Dancy II TE. 2011. A qualitative investigation of factors promoting the retention and persistence of students of color in STEM. J Negro Educ 80:491–504.
42. Dika SL, D’Amico MM. 2016. Early experiences and integration in the persistence of first-generation college students in STEM and non-STEM majors. J Res Sci Teach 53:368–383. https://doi.org/10.1002/tea.21301
43. Stephens NM, Fryberg SA, Markus HR, Johnson CS, Covarrubias R. 2012. Unseen disadvantage: how American universities’ focus on independence undermines the academic performance of first-generation college students. J Pers Soc Psychol 102:1178–1197. https://doi.org/10.1037/a0027143
44. Knekta E, Runyon C, Eddy S, Brickman P. 2019. One size doesn't fit all: using factor analysis to gather validity evidence when using surveys in your research. CBE Life Sci Educ 18:rm1. https://doi.org/10.1187/cbe.18-04-0064