Understanding Successful Engagement of Women and Underrepresented Students in a Multidisciplinary Course-Based Undergraduate Research Experience (CURE)

Lisa Dierker[1], Janet Rosenbaum[2], Jeremiah Piña[3], Ajíšafé Adaján[4]
[1]Wesleyan University, [2]SUNY, Downstate Medical Center,
[3]Smith College, [4]Creative Research Solutions

ABSTRACT
Course-based undergraduate research experiences (CURE) provide valuable opportunities to large numbers of students relatively early in their academic careers and have the potential to attract and retain women and students from underrepresented minority groups both in the sciences and in other technical fields requiring quantitative research literacy. To evaluate the relative success of a multidisciplinary CURE, we compared background characteristics, course experiences and outcomes of men and women under-represented (URM) and non-underrepresented students. Though URM and non-URM students of both genders differed on many background characteristics, and self-reported course experiences, with few exceptions, positive course outcomes and predictors of those outcomes did not differ by URM/gender group. The Passion-Driven Statistics CURE aims to equip the future STEM workforce with the data analysis skills and reasoning needed across industries. Additional research is needed to determine whether this CURE may influence educational and career trajectories for women and URM students.

KEYWORD: CURE, multidisciplinary, higher education.

INTRODUCTION

Decades of scholarship has supported the value of undergraduate research experiences in attracting and retaining women and underrepresented minority students in the sciences (American Association for the Advancement of Science, 2011; Auchincloss, et al., 2014) and in technical fields requiring quantitative research literacy (U.S. National Science Foundation, 2021). Course-based undergraduate research experiences (CURE) have been shown to be particularly promising given their ability to provide research opportunities to large numbers of students and to engage them in research relatively early in their academic careers (e.g., Rowland et al., 2012; Harrison et al., 2011). This gives CURE’s the potential to exert a greater influence on students’ academic and career paths than research internships that typically occur late in an undergraduate’s academic program and after many students have exited the major or withdrawn from college (Hunter et al., 2007; Auchincloss, et al., 2014). Studies comparing CURE instruction with research internships and laboratory learning experiences have generally found that students report many of the same gains including a sense of ownership of the science projects (Hanauer et al., 2012; Hanauer and Dolan, 2014) and higher levels of persistence in science or medicine (Hanauer et al., 2012). To date however, few studies have explored how different students may experience CUREs or realize different outcomes (Auchincloss, et al., 2014).

In previous publications, we have described the development of Passion-Driven Statistics, a multidisciplinary CURE aimed at engaging students in course-based research across both divisional and departmental boundaries (Dierker, et al., 2012). Funded by the U.S. National Science Foundation and closely following the recommendations outlined in the Course-Based Undergraduate Research Experiences Network report (Auchincloss, et al., 2014) the curriculum is designed around authentic research projects of student's own choosing and offers individualized hands-on experience in quantitative research and applied statistics through engagement with real world data and statistical software. It has been implemented within statistics courses, research methods courses, data science courses, and mentored research experiences with students from a wide range of academic settings. Liberal arts colleges, large state universities, regional colleges/universities, medical schools, community colleges, and high schools have all successfully implemented the model (Dierker, et al., 2018a).

Based on the student's choice of data, each generates testable hypotheses for their chosen dataset; conducts a literature review on their topic of interest; works to refine or broaden their research questions based on information they collect; prepares data for analysis; selects and conducts descriptive and inferential statistical analyses; and evaluates, interprets, and presents research findings. These activities are not presented or experienced as distinct stages but rather, as a series of ongoing, interactive tasks. Learning materials and teaching strategies were designed to be structured enough to allow students to consistently move forward with their research projects, yet broad enough to encourage them to explore their questions creatively and independently, letting students actively drive the decisions involved in inquiry. In this way, the support each student receives is dictated by their own research question and the results at each stage of their project. Students work to become fluent in the use of common scientific practices and to understand the iterative nature of research. They engage in collaboration with peers and instructors in the production of relevant work that builds on current scientific knowledge, and they share their novel discoveries with their local and academic communities (Auchincloss, et al., 2014)

Research evaluating the Passion-Driven Statistics CURE has been promising. Based on data collected at the liberal arts college where the curriculum was originally developed, we demonstrated that the CURE enrolls larger numbers of under-represented minority (URM) students compared to a traditional introductory statistics curriculum (Dierker, et al., 2015). Further, while URM students considered the material in the CURE more difficult than non-URM students, they demonstrated similar levels of increased confidence in applied skills and interest in follow up courses. URM students were also found to be twice as likely as non-URM students to report that their interest in conducting research increased after completing the project-based course (Dierker et al., 2016). Compared to students enrolled in a traditional introductory statistics course, those completing the Passion-Driven Statistics CURE have been shown to be significantly more likely to report an increase in confidence between a pre and post course survey with regard to choosing the correct statistical test, managing data, and writing syntax or code to run statistical analyses. CURE students were also more likely than students in the traditional statistics course to show an increase in their interest in pursuing advanced course work in research and applied statistics (Dierker et al., 2018b). Because of the curriculum's focus on programming in the context of quantitative research, we also compared enrollment in the Passion-Driven Statistics CURE to traditional introductory programming experiences, revealing higher rates of women and URM enrollment compared to both a general introductory programming course and an introductory course representing a gate to the computer science major (Cooper & Dierker, 2017). The Passion-Driven Statistics CURE has also been shown to impact students' academic trajectory. Using causal inference techniques to achieve matched comparisons across three different statistics courses, students originally enrolled in the Passion-Driven Statistics CURE were significantly more likely to take at least one additional undergraduate course focused on research methods, statistical concepts, applied data analysis, and/or use of statistical software compared to students taking either a psychology statistics course or math statistics course (Nazzaro, et al., 2020).

Based on survey data from a large sample of students enrolled in the Passion-Driven Statistics CURE across a range of undergraduate courses and schools in the United States, the present paper compares background characteristics, learning experiences and course outcomes for male and female URM and non-URM students to determine whether there are systematic differences by gender and/or ethnic minority status.

Methods

Participants

Pre and post course survey data from 2125 students enrolled in a CURE using the Passion-Driven Statistics curriculum between fall of 2018 and spring of 2022 at one of 38 post-secondary institutions were examined. Participants included 1316 (61.9%) students from liberal arts colleges, 176 (8.3%) students from flagship state universities, 501 (23.6%) from public regional colleges/universities and 132 (6.2%) from community colleges. The sample included 571 (26.9%) students who self-identified as being from an underrepresented group (URM: Hispanic, n=372, African American n=189, or Native American, n=10) and 1554 (73.1%) students who self-identified as not underrepresented (non-URM). Students were also asked to self-report their gender: 762 (35.9%) students identified as men and 1363 (64.1%) students identified as women.

CURE Features

At each school, the CURE was delivered in the context of research projects of students' own choosing. The curriculum focused on the decisions and skills involved in asking and answering questions with data with the goal of telling accurate and engaging stories. To accomplish this, students were provided with one or more large data sets and corresponding data dictionaries. Based on the student's choice of data, each generated testable hypotheses, prepared data for analysis, selected and used descriptive and inferential statistical tools; and evaluated, interpreted, and presented research findings, orally, graphically, and/or as text. Final projects most often included a scientific poster presentation allowing students to communicate methods, results, and insights. In the absence of a poster presentation, final projects included electronic submission of scientific posters, final lab reports, and/or research proposals. The CURE included statistics courses =, discipline specific research methods courses, and data science/analysis courses.

Each CURE was semester-long, meeting 1, 2, or 3 times each week. Each instructor chose the statistical software platform to be used in their course. These included R, SAS, Stata, JMP, SPSS, or StatCrunch. At each school, the project-based course typically counted toward or was required by one or more majors including sociology, psychology, political science, human development, criminal justice, business, nursing, biology, and/or neuroscience and behavior.

Measures

Data were drawn from student surveys completed before (pre) and after (post) each course. The pre course survey was completed prior to the end of the second week of classes and the post course survey was completed during the last weeks of the semester. Each survey took approximately 10-15 minutes to complete.

Demographics and academic background characteristics

The pre course survey included questions assessing age, gender, and race/ethnicity (coded as dichotomously as underrepresented – Hispanic, Black, and/or Native American vs. all other race/ethnicities). Participants also reported whether they were the first generation in their family to attend college or were eligible for free/reduced lunches during high school. Interest in the course was assessed based on the question how likely you would be to enroll in this course if it were not required from 1 (very unlikely) to 5 (very likely).

Academic background was assessed by questions about having taken a previous statistics course (high school or college), any prior programming experience, e.g., R, SAS, Stata, Python, C++, Java, HTML, etc. (yes/no) and self-confidence when it comes to learning programming from 1 (strongly disagree) to 5 (strongly agree). Self-perceived skills in mathematics were measured by the question, —How good at mathematics are you? on a scale from 1 (very poor) to 5 (very good). Anxiety about the course was also assessed (—The thought of being enrolled in this course makes me nervous) from 1 (strongly disagree) to 5 (strongly agree) (Wise, 1985). Class standing was recorded as first/second year undergraduate vs. third/fourth year undergraduate

Items from the Classroom Undergraduate Research Experience Survey (CURE) were included in the pre course survey to measure *prior experience with research*. Areas of experience included conducting a quantitative project entirely of your own design, reading primary source scientific literature, collecting data, analyzing data, and presenting a poster at a science or research poster session. Response options ranged from 1 (no experience or feel inexperienced) to 5 (extensive experience or mastered this element). A mean score was calculated as an aggregate measure of previous research experience (Lopatto, 2008).

Student Experiences in the Course

A variable was created to indicate whether the course was taken on-line during the COVID pandemic (spring 2019 through summer 2021). Since each instructor selected and supported the software used in their course, a variable was also created to indicate whether students learned a code-based software program (i.e., R, SAS, Stata, python) or a point-and-click interface (i.e., JMP, SPSS, JAMOVI or StatCrunch).

Items on the satisfaction scale from the Undergraduate Research Student Self-Assessment (URSSA) were also administered (Hunter, et al., 2009; Weston & Laursen, 2015). Students rated their working relationship and amount of time spent with their instructor, teaching apprentice or peer mentor, and research group members. Response options ranged from 1 (poor) to 5 (excellent) and were averaged to create an aggregate measure of support.

Students' perceptions of course rigor was measured by questions including —How challenging did you find this course? from 1 (not at all) to 5 (the most challenging), —How hard did you work in this course from 1 (not at all hard) to 5 (extremely hard) and —Was this course more challenging, less challenging or similarly challenging compared to other college courses you have taken (dichotomized as more challenging vs. similarly or less challenging)?

Overall impressions of the success of the course were measured with the question —Did you accomplish more than you expected, less than you expected or about the same as you expected (dichotomized as more vs. less or the same)? and —Did you find this course more useful, less useful or similarly useful compared to other college courses you have taken (dichotomized as more useful vs. similarly or less useful)?

Questions addressing *student engagement* were drawn from the survey used by Gasiewski et al., (2012). Areas assessed included: asking questions in class, discussing course grades or assignments with the instructor, attending the instructor's office hours, tutoring other students in the course, preparing for class sessions by completing assigned readings and/or videos, reviewing class materials before they were covered, attending review or help sessions to enhance understanding of course content, studying with students from the course, and participating in class discussions. Response options ranged from 1 (never) to 5 (always). Items were averaged to create an aggregate measure of students' engagement.

Outcomes

Core scales from the URSSA were used to measure self-reported gains in thinking and working like a scientist (e.g., analyzing data for patterns, identifying limitations of research methods), personal gains related to research (e.g., ability to contribute to science, comfort discussing scientific concepts, patience in the slow pace of research), gains in research skills (e.g., preparing a scientific poster, keeping a detailed lab notebook) and positive attitudes and behaviors as a researcher (e.g. feeling like a scientist, feeling responsible for the project, feeling part of a scientific community). Previous research has shown that the URSSA represents separate but related constructs from the undergraduate research experience. Average scores form reliable, moderate to highly correlated composite measures. URSSA scales have been shown to correlate with ratings of satisfaction with external aspects of the research experience (Hunter, et al., 2009; Weston & Laursen, 2015). Response options ranged from 1 (none) to 5 (a great deal). Items were averaged to create aggregate measures for each scale. Composite variables exceeded accepted standards for reliability (i.e., alpha >.0.80)

Students were asked additional questions in both the pre and post course surveys pertaining to their interest in conducting research from 1 (not at all interested) to 5 (extremely interested) and likelihood/intention of (a) pursuing advanced coursework in statistics or data analysis (b) taking any statistics or data analysis courses in the future from 1 (definitely not) to 5 (definitely yes), (c) using statistics as they complete the remainder of their degree program and (d) using statistics in the field in which they hope to be employed when they finish school from 1 (never) to 5 (frequently). Responses on the pre survey were subtracted from responses to the same post survey questions and dichotomous variables were created and averaged to indicate the number of positive *increases in interest, or intention* (range 0 to 5).

Analyses

Chi-square Tests of Independence and ANOVA were used to evaluate differences by URM/gender group for categorical and continuous variables, respectively. A Bonferroni adjustment was used to evaluate post hoc paired comparisons for Chi Square Tests of Independence and the Duncan test was used to evaluate post hoc paired comparisons for ANOVA. Next, multiple linear regression analysis was used to explain variation in each outcome measure. URM status, gender, a URM x gender interaction, individual background characteristics, and course experience were used as predictor in each regression model. The results are shown as standardized coefficients (beta). Cases with missing data were deleted listwise. Interactions between significant main effects and URM/gender grouping were also tested. Those that were significant are reported as pos hoc findings.

Results

Demographic and Academic Background

A comparison of demographic and academic background characteristics by race/gender group is presented in Table 1. Both gender and URM groups were similar in terms of level of prior research experience and having previously taken a statistics course in high school or college. URM women, however, were older than non-URM women and both men groups of men. Further, URM men were significantly older than non-URM students of both genders. URM students of both genders were significantly less likely to have taken the course in their first or second year of college, more likely to be attending a public college or university, to be a first-generation college student and to have been eligible for free or reduced lunches in high school compared to non-URM students of both genders. URM women were also significantly more likely to have been eligible for free or reduced lunches in high school compared to URM men.

Overall, men self-reported being better at mathematics than women. Further, non-URM women self-reported being better at mathematics than URM women. URM women reported higher levels of nervousness about learning statistics than non-URM students of both genders. Nervousness reported by non-URM women was also significantly greater than reports by both URM and non-URM men. Overall, men were more likely to have had previous programming experience and were more likely to have taken the course if it were not required compared to women.

Table 1. Demographic and academic background characteristics of students taking the research course by URM[1] and gender

	URM women n = 404	non-URM women n = 959	URM men n = 167	non-URM men n = 595	Statistics[2]
	n (%)	n (%)	n (%)	n (%)	
1st or 2nd year student	165 (40.8%)[b]	511 (53.3%)[a]	60 (35.9%)[b]	315 (52.9%)[a]	$X^2(3) = 32.7$, $p < .0001$
Private college/university vs. public	127 (31.4%)[b]	671 (70.0%)[a]	76 (45.5%)[b]	442 (74.3%)[a]	$X^2(3) = 243.3$, $p < .001$
First generation college student	192 (48.7%)[a]	152 (16.1%)[b]	74 (45.4%)[a]	98 (16.8%)[b]	$X^2(3) = 217.1$, $p < .001$
Eligible for free/reduced lunch in high school	234 (60.9%)[a]	137 (15.1%)[c]	75 (48.1%)[b]	91 (16.1%)[c]	$X^2(3) = 360.9$, $p < .001$
Prior statistics experience[3]	241 (59.7%)	496 (51.7%)	96 (57.5%)	336 (56.5%)	*n.s.*
Programming experience	86 (21.3%)[b]	211 (22.1%)[b]	58 (34.9%)[a]	259 (43.6%)[a]	$X^2(3) = 98.4$, $p < .001$
	Mean (s.d.)	Mean (s.d.)	Mean (s.d.)	Mean (s.d.)	
Student age	23.6[a] (6.88)	20.8[c] (4.28)	22.8[b] (6.55)	20.7[c] (2.92)	$F(3, 2100) = 43.5$, $p < .0001$
CURE prior research experience subscale	2.5[b] (0.87)	2.6 [a,b] (0.74)	2.6 [a,b] (0.79)	2.7[a] (0.76)	*n.s.*

How good at mathematics are you? (1 = v poor to 5 = v good)	3.0[c] (0.88)	3.3[b] (0.89)	3.4[a] (0.80)	3.4[a] (0.87)	$F(3, 2116) = 14.3$, $p < .0001$
I have a lot of self-confidence when it comes to learning programming (1 = str disagree to 5 = str agree)	2.7[b] (0.94)	2.3[c] (0.95)	2.9[a] (0.94)	2.8[a,b] (1.06)	$F(3, 2116) = 31.6$, $p < .0001$
The thought of being enrolled in a statistics course makes me nervous (1 = str disagree to 5 = str agree)	2.6[a] (1.19)	2.4[b] (1.08)	3.0[c] (1.14)	2.9[c] (1.09)	$F(3, 2117) = 46.6$, $p < .0001$
How likely you would take a course in statistics if not required? (1 = v unlikely to 5 = v likely)	2.4[b] (1.18)	2.6[b] (1.20)	3.0[a] (1.22)	3.1[a] (1.12)	$F(3, 2116) = 31.5$, $p < .0001$

[1]Under-represented students included those self-identifying as Black, Hispanic, or Native American.
[2]Percentages are based on the number of respondents completing each item or scale.
[3]Aggregate of responses querying prior stats experience in high school, high school AP/IB, or college

Course Experiences

A comparison of course experiences by race/gender group is presented in Table 2. Significantly more URM women took the research course during the COVID pandemic compared to URM and non-URM men. Further, a larger proportion of URM women were enrolled in a course using code-based software rather than a point and click interface when compared to non-URM men and women.

Non-URM men were significantly less likely to rate the course as very or extremely challenging compared to women and URM men. They were also less likely than each of the comparison groups to report working very or extremely hard in the course. Non-URM women were also significantly less likely to rate the course as very or extremely challenging compared to URM students of both genders. Non-URM women were less likely to report working very or extremely hard in the course, but only compared to URM women.

Both URM and non-URM women were more likely than non-URM men to report that they had accomplished more than they had expected in the course. In contrast, non-URM men reported higher levels of engagement in the course compared to both URM and non-URM women. They were also more likely to rate the course as more useful than other college courses but only compared to URM women.

Table 2. Course experiences of students taking the research course by URM[1] and gender

	URM women n = 404	non-URM women n = 959	URM men n = 167	non-URM men n = 595	Statistics[2]
	n (%)	n (%)	n (%)	n (%)	
Course Format					
Course taken on-line during COVID	196 (48.5%)[a]	395 (41.2%)[a,b]	55 (32.9%)[b]	225 (37.8%)[b]	$X^2(3) = 16.4$, $p < .0009$
Course used code-based software	345 (85.4%)[a]	708 (73.8%)[b]	132 (79.0%)[a,b]	429 (72.1%)[b]	$X^2(3) = 27.8$, $p < .0001$
Rigor	n (%)	n (%)	n (%)	n (%)	
Found course very (4) or extremely (5) challenging	189 (47.4%)[a]	293 (30.9%)[b]	74 (45.7%)[a]	123 (20.9%)[c]	$X^2(3) =90.2$, $p < .0001$
Worked very (4) or extremely (5) hard in course	276 (69.2%)[a]	518 (54.6%)[b]	101 (62.4%)[a,b]	230 (39.1%)[c]	$X^2(3) =93.9$, $p < .0001$
Success					
Accomplished more than expected	226 (56.6%)[a]	484 (51.2%)[a]	77 (47.5%)[a,b]	236 (40.6%)[b]	$X^2(3) = 28.0$, $p < .0001$
Course was more useful than other college courses	169 (43.1%)[b]	456 (48.6%)[a,b]	77 (49.4%)[a,b]	317 (54.6%)[a]	$X^2(3) = 12.6$, $p = .0055$
Support	Mean (s.d.)	Mean (s.d.)	Mean (s.d.)	Mean (s.d.)	
Engagement with the course	2.9[a] (0.72)	2.9[a] (0.65)	2.8[ab] (0.68)	2.8[b] (0.69)	$F(3, 2096) = 5.05$, $p < .0017$
URSSA instructional support[5]	3.3 (0.60)	3.3 (0.54)	3.2 (0.56)	3.2 (0.56)	*n.s.*

[1]Under-represented students included those self-identifying as Black, Hispanic, or Native American.
[2]Percentages are based on the number of respondents completing each item or

scale.

Course Outcomes

A comparison of course outcomes by gender/URM group is presented in Table 3. Gains in thinking and working like a scientist (e.g., analyzing data for patterns, identifying limitations of research methods), gains related to research (e.g., ability to contribute to science, comfort discussing scientific concepts, patience in the slow pace of research), and positive research behaviors (e.g., feeling like a scientist, feeling responsible for the project, feeling part of a scientific community) did not differ across URM/gender group. Increased interest and expectations around research and using statistics were also found to be similar. However, average gains in research skills were rated higher among URM and non-URM women compared to non-URM men.

Table 3. Course outcomes of students taking the research course by URM[1] and gender

	URM[1] women n = 404	Non-URM women n = 959	URM men n = 167	Non-URM men n = 595	Statistics[2]
	n (%)	n (%)	n (%)	n (%)	
Increased interest pre to post[3]					
Increased interest in conducting research	136 (33.7%)	315 (32.9%)	107 (35.9%)	194 (32.6%)	*n.s.*
Increased interest in pursuing advanced coursework in stats or data science	126 (31.2%)	282 (29.4%)	43 (25.8%)	186 (31.3%)	*n.s.*
Increase in intention to take any more stats or data science courses	121 (30.0%)	259 (27.0%)	41 (24.6%)	170 (28.6%)	*n.s.*
Increased expectation pre to post[3]					
Increased expectation of using statistics in future coursework	103 (25.5%)	191 (19.9%)	48 (28.7%)	145 (24.4%)	*n.s.*
Increased expectation of using statistics in future employment	101 (25.0%)	228 (23.8%)	40 (24.0%)	157 (26.4%)	*n.s.*
	Mean (s.d.)	Mean (s.d.)	Mean (s.d.)	Mean (s.d.)	
Proportion showing an increase in interest/expectation[4]	0.32 (0.31)	0.29 (0.30)	0.32 (0.30)	0.33 (0.31)	*n.s.*

Post responses to URSSA subscales (1 = no gain to 5 = great gain)					
Thinking like a scientist	4.0 (0.83)	4.1 (0.87)	4.0 (0.89)	4.0 (0.85)	*n.s.*
Gains related to research	3.8 (0.92)	3.8 (0.95)	3.7 (0.95)	3.7 (1.00)	*n.s.*
Gains in research skills	3.7[a] (0.90)	3.7[a] (0.90)	3.6[a,b] (0.90)	3.5[b] (0.96)	$F(3, 2058) = 6.1$, $p < .0004$
Positive attitudes and behaviors	3.4 (0.99)	3.3 (0.95)	3.4 (0.93)	3.3 (0.99)	*n.s.*

[1]Under-represented students included those self-identifying as Black, Hispanic, or Native American
[2]Percentages are based on the number of respondents completing the item or scale
[3]Counts are based on the number of respondents with an increase in response scores across pre/post items
[4]Proportion of responses indicating increase among the five pre/post items

Next, multivariate regression models were run to evaluate whether student background and course experiences predict positive course outcomes. Table 4 presents model estimates for each main effect as well as the URM x gender interaction. Prior research experience, level of engagement in the course, satisfaction with instructional support, perceiving the course as less challenging, accomplishing more than expected, and feeling the course was more useful than other courses were each independently associated with each course outcome. In addition, higher levels of hard work in the course and self-confidence in learning to program were associated with each course outcome, except for increased interest and expectations pre/post around research. A higher likelihood of taking the course if not required was associated with gains related to research and positive research behaviors. The remaining effects were significant for unique course outcomes. Higher levels of nervousness about being enrolled in the course for example was associated with gains in positive research behaviors, first- or second-year status was associated with gains in research skills and experience in a prior statistics course and self-reported skill in mathematics was associated with increased interest and expectations around research and statistics.

Interactions between significant main effects and URM/gender grouping were also tested. Most interactions were non-significant demonstrating that predictors of course outcomes were consistent across URM/gender groups. However, there was a significant interaction between URM/gender group and how challenging the course was judged to be and how hard students worked for each of the course outcomes measured by the URSSA subscales. Models estimated for individual URM/gender subgroups revealed that rating the course as less challenging predicted each course outcomes for women, both URM and non-URM, but not men. In contrast, higher levels of hard work predicted gains in research skills and positive research behaviors for non-URM women, and men (both URM and non-URM), but not URM women. Higher levels of hard work predicted gains in thinking and working like a scientist and personal gains related to research for non-URM students (both men and women) but not URM students.

Table 4. Multiple regression models predicting each course outcome (beta weight).

	Future interest/ expect.	Thinking like a scientist (URSSA)	Research gains (URSSA)	Research skills (URSSA)	Positive attitudes / behaviors (URSSA)
URM	0.01	0.06	0.03	0.11	0.02
Gender	0.03	-0.05	-0.08	**-0.12****	-0.02
URM x Gender	0.01	-0.03	-0.03	-0.07	0.07
Student age	0.00	0.00	-0.01	-0.01	0.00
1st or 2nd year student	0.00	0.04	0.05	**0.09***	0.02
Private college/university vs. public	-0.02	0.05	-0.04	0.00	0.01
First generation college student	0.01	0.02	0.08	0.03	0.03
Eligible for free/reduced lunch in high school	0.03	0.02	0.05	0.06	0.11
Prior statistics experience	**0.03***	0.00	0.06	0.07	0.04
Programming experience	0.03	-0.02	-0.06	-0.05	-0.01
CURE prior research experience subscale	**-0.02***	**0.10*****	**0.13*****	**0.07****	**0.13*****
How good at mathematics are you?	**-0.02***	0.00	0.04	0.00	0.02
I have a lot of self-confidence when it	0.00	**0.05****	**0.07****	**0.08*****	**0.09*****

comes to learning programming					
The thought of being enrolled in a statistics course makes me nervous	-0.01	0.02	0.03	0.03	**0.04***
How likely you would take a course in statistics if not required?	0.01	0.02	**0.04***	0.03	**0.09*****
On-line during COVID	0.00	0.04	0.01	0.01	0.02
Course used code-based software	0.03	-0.03	-0.05	0.03	0.04
Challenging	**-0.05*****	**-0.15*****	**-0.15*****	**-0.14*****	**-0.10****
Worked Hard	0.01	**0.11*****	**0.12*****	**0.14*****	**0.14*****
Accomplished more than expected	**0.07*****	**0.40*****	**0.42*****	**0.34*****	**0.31*****
More useful than other college courses	**0.13*****	**0.34*****	**0.29*****	**0.24*****	**0.19*****
Engagement with the course	**0.06*****	**0.11*****	**0.17*****	**0.21*****	**0.28*****
URSSA instructional support	**0.03***	**0.36*****	**0.44*****	**0.47*****	**0.45*****

*p<.05; ** p<.01; *** p<.001

Discussion

We have described in previous reports the development of the Passion-Driven Statistics CURE, its success in attracting higher rates of URM students compared to a traditional math statistics curriculum (Dierker et al., 2012; Dierker et al., 2015) and the role of the CURE in influencing future course choices (Nazzaro, et al., 2020). Based on survey data from a large sample of students enrolling in the CURE across diverse undergraduate settings, the present study compared men and women URM and non-URM students on background characteristics, learning experiences and course outcomes to evaluate factors that increase confidence and foster positive attitudes toward research and applied data analysis opportunities. Notably, URM students were older than non-URM students, less likely to have taken the CURE in their first or second year of college, more likely to be attending a public college or university, to be a first-generation college student and to have been eligible for free or reduced lunches in high school, while women self-reported less aptitude in mathematics, less interest in the course if it were not required, more nervousness about taking the course and being less likely to have prior programming experience.

Despite these many pre course differences, none were uniquely related to course outcomes for individual URM/gender groups. Instead, confidence in learning to program, though higher for men than women, was found to be associated with self-reported gains in thinking like a scientist, research behaviors, research skills, and attitudes and behaviors for each of the URM/gender groups. Lower levels of programming confidence and lower ratings of mathematics abilities among women, despite similar performance compared to men has been well documented (Liberatore & Wagner, 2022; Baird & Keene, 2019). Given that higher confidence in ability to learn programming predicted positive course outcomes even after controlling for prior programming and individual experiences within the course, this CURE would benefit from additional practices that work to overcome these negative cultural beliefs. For example, Dweck's (2016) research finds that emphasizing math, science, and programming abilities as a learnable skill rather than an unchangeable trait can impact a variety of educational and career-relevant outcomes.

The background characteristic found to be similar across URM/gender groups and also a consistent predictor of positive course outcomes was prior research experience. Because URM students were less likely to enroll in the CURE in the first two years of their undergraduate studies, it seems that comparable research experience may not have been reached by URM and non-URM students at comparable educational stages. Given that early participation in CUREs has the best chance of influencing students' selection of both future courses as well as a major discipline, these findings confirm the need to provide CURE opportunities to URM students early and often in their academic careers to help to reverse the 'leaky pipeline' in which large numbers of URM students turn away from courses focused on science and technology (Linnenbrink-Garcia et al., 2018).

Notably, most experiences with the CURE differed by URM/gender group but were found to be consistently associated with positive course outcomes irrespective of URM and gender status. These included a higher level of engagement, accomplishing more than expected in the course, finding the course more useful than other college courses, and satisfaction with instructional support. How hard students worked in the course and how challenging they found the CURE however differed by URM/gender group, and though URM students reported working harder in the course and finding the course more challenging on average, it was finding the CURE less challenging that predicted positive outcomes for women but not men and working harder in the CURE that predicted gains most often for non-URM students rather than URM students. Garfield and Ben-Zvi (2007) stress the benefit to students of struggling to solve conceptually difficult problems and associating these problems with other areas of relevant knowledge. The productive resolution of temporary intellectual roadblocks is a common occurrence in this and other CUREs. Differences found in predicting positive outcomes for women and URM students may highlight the need to present challenges that allow students to resolve moments of intellectual struggle so that they can experience a stronger sense of accomplishment about their work despite the difficulty. An important ingredient for the kind of productive struggle that promotes a sense of accomplishment is individualized support (Pasquale, 2015) which was also found to independently predict positive course outcomes in the present analyses.

Finally, it is important to note that average gains in research skills were rated higher among women (both URM and non-URM) compared to non-URM men. This finding was replicated in multivariate models controlling for students' background and course experiences. Previous work has suggested that research experiences may be particularly useful for women and underrepresented minority students given the opportunities to develop relationships with scientists and peers (Barlow and Villarejo, 2004; Eagan et al., 2011). Although additional research is needed to determine the mechanisms that may contribute to observed differences in self-reported research skills, women and URM students will benefit the most if CUREs are widely available early in their education. Because men are more likely to self-reported less interest in the CURE if it were not required, these opportunities need to be standard within the college experience.

Though course outcomes were self-reported rather than based on standardized exams, previous research on project-based learning has shown that it is important to examine engagement as well as additional experiences and outcomes beyond typical content knowledge (Dochy, et al., 2003). It is these additional outcomes and beliefs that are most important to continued persistence in quantitative courses (Wigfield & Eccles, 2000). Research on factors that lead to retention in science majors for example indicates that increased science identity (Estrada et al., 2011), ability to navigate uncertainty, and resilience are important precursors to a sense of belonging and ultimate retention (Maton and Hrabowski, 2004; Seymour et al., 2004). In fact, previous scholarship has shown that students' interest in research is an important factor in increasing skills and a sense of identity as a researcher (Graham, et al., 2013; Hurtado, et al., 2009). While the present data was drawn from undergraduate students in the United States, we have previously reported on the use of the Passion-Driven Statistics model in Ghana, West Africa (Awuah, et al., 2020). Additional research is needed to determine whether this CURE may influence educational and career trajectories specifically for women and URM students.

Taken together, the present findings provide important insights into the multifaceted approach that will help to increase engagement of women and URM students in research-focused disciplines. First, they should be provided with early and frequent exposure to research experiences. Second, individualized support and mentorship should be offered. Project-based curricula present students with intellectually challenging problems, providing support to help them overcome these challenges, and fostering a sense of accomplishment and engagement. URM students and women should also be supported to develop a strong science identity by highlighting the relevance of their work and the impact it can have on their communities and society as a whole. Continuous evaluation of the effectiveness of curriculum in engaging URM students and women is also critical. By implementing these suggestions, educators can help increase engagement among women and URM students, ultimately leading to improved outcomes and increased retention in STEM fields.

References

American Association for the Advancement of Science. (2011). *Vision and Change in Undergraduate Biology Education*. https://visionandchange.org/

Auchincloss, L. C., Laursen, S. L., Branchaw, J. L., Eagan, K., Graham, M., Hanauer, D. I., Lawrie, G., McLinn, C. M., Pelaez, N., Rowland, S., Towns, M., Trautmann, N. M., Varma-Nelson, P., Weston, T. J., & Dolan, E. L. (2014). Assessment of course-based undergraduate research experiences: A meeting report. *CBE Life Sciences Education*, *13*(1), 29–40. https://doi.org/10.1187/cbe.14-01-0004

Awuah, R., Gallagher, K. M., & Dierker, L. C. (2020). Taking Project-Based Statistics Abroad: Learning experiences and outcomes of a project-based statistics course in West Africa. *Statistics Education Research Journal*, *19*(3), 18–31.

Baird, C. L., & Keene, J. R. (2019). Closing the Gender Gap in Math Confidence: Gender and Race/Ethnic Similarities and Differences. *International Journal of Gender, Science and Technology*, *10*(3), 378–410. https://genderandset.open.ac.uk/index.php/genderandset/article/view/452

Barlow, A. & Villarejo, M. (2004). Making a difference for minorities: Evaluation of an educational enrichment program. *Journal of Research in Science Teaching*, *41*(9), 861–881. https://doi.org/10.1002/tea.20029

Cooper, J. L., & Dierker, L. C. (2017). Increasing exposure to programming: A comparison of demographic characteristics of students enrolled in introductory computer science programming courses vs. A multidisciplinary data analysis course. *International Research in Higher Education*, *2*(1), 92–100. https://doi.org/DOI: 10.5430/irhe.v2n1p92

Dierker, L. C., Alexander, J., Cooper, J. L., Selya, A., Rose, J., & Dasgupta, N. (2016). Engaging Diverse Students in Statistical Inquiry: A Comparison of Learning Experiences and Outcomes of Under-Represented and Non-Underrepresented Students Enrolled in a Multidisciplinary Project-Based Statistics Course. *International Journal for the Scholarship of Teaching and Learning*, *10*(1). https://doi.org/10.20429/ijsotl.2016.100102

Dierker, L. C., Cooper, J. L., Selya, A., Alexander, J., & Rose, J. (2015). Evaluating Access: Comparing enrollment patterns in traditional versus multidisciplinary, project-based introductory statistics courses. *Journal of Interdisciplinary Studies in Education*, *4*(1), 22–37. https://www.ojed.org/index.php/jise/article/view/1599

Dierker, L. C., Robertson Evia, J., Singer-Freeman, K., Woods, K., Zupkus, J., Arnholt, A., Moliski, E. G., Deckard, N. D., Gallagher, K., & Rose, J. (2018a). Project-Based Learning in Introductory Statistics: Comparing Course Experiences and Predicting Positive Outcomes for Students from Diverse

Educational Settings. *International Journal of Educational Technology and Learning*, *3*(2), 52–64. https://doi.org/DOI:10.20448/2003.32.52.64

Dierker, L. C., Flaming, K., Cooper, J. L., Singer-Freeman, K., Germano, K., & Rose, J. (2018b). Evaluating Impact: A Comparison of Learning Experiences and Outcomes of Students Completing A Traditional Versus Multidisciplinary, Project-Based Introductory Statistics Course. *International Journal of Education, Training and Learning*, *2*(16–28). https://doi.org/10.33094/6.2017.2018.21.16.28

Dierker, L. C., Kaparakis, E., Rose, J., Selya, A., & Beveridge, D. (2012). Strength in Numbers: A Multidisciplinary, Project-Based Course in Introductory Statistics. *Journal of Effective Teaching*, *12*(2), 4–14. https://eric.ed.gov/?id=EJ1092198

Dochy, F., Segers, M., Van den Bossche, P., & Gijbels, D. (2003). Effects of problem-based learning: A meta-analysis. *Learning and Instruction*, *13*(5), 533–568. https://doi.org/10.1016/S0959-4752(02)00025-7

Dweck, C. S. (2016). *Mindset*. Random House.

Eagan, M. K., Sharkness, J., Hurtado, S., Hurtado, S., Mosqueda, C. M., & Chang, M. J. (2011). Engaging Undergraduates in Science Research: Not Just About Faculty Willingness. *Res High Educ*, *52*, 151–177. https://doi.org/10.1007/s11162-010-9189-9

Estrada, M., Woodcock, A., Hernandez, P. R., & Schultz, P. (2011). Toward a model of social influence that explains minority student integration into the scientific community. *J Educ Psychol*, *103*, 206–222. https://pubmed.ncbi.nlm.nih.gov/24591501/

Garfield, J., & Ben-Zvi, D. (2007). How Students Learn Statistics Revisited: A Current Review of Research on Teaching and Learning Statistics. *International Statistical Review*, *75*(3), 372–396. https://doi.org/10.1111/j.1751-5823.2007.00029.x

Gasiewski, J. A., Eagan, M. K., Garcia, G. A., Hurtado, S., & Change, M. J. (2012). From gatekeeping to engagement: A multi-contextual, mixed method study of student academic engagement in introductory STEM courses. *Res High Educ*, *53*, 229–261.

Graham, M. J., Frederick, J., Byars-Winston, A., Hunter, A.-B., & Handelsman, J. (2013). Increasing persistence of college students in STEM. Science. *Science*, *341*, 1455–1456. https://pubmed.ncbi.nlm.nih.gov/24072909/

Hanauer, D. I., & Dolan, E. L. (2014). The project ownership survey: Measuring differences in scientific inquiry experiences. *CBE Life Sciences Education*, *13*(1), 149–158. https://doi.org/10.1187/cbe.13-06-0123

Hanauer, D. I., Frederick, J., Fotinakes, B., & Strobel, S. A. (2012). Linguistic analysis of project ownership for undergraduate research experiences. *CBE Life Sciences Education*, *11*(4), 378–385. https://doi.org/10.1187/cbe.12-04-0043

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 Science Education*, *10*, 279-286.

Hunter, A. B., Laursen, S. L., & Seymour, E. (2007). Becoming a scientist: The role of undergraduate research in students' cognitive, personal, and professional development. *Science Education*, *91*(1), 36–74

Hunter, A. B., Weston, T. J., Laursen, S. L., & Thiry, H. (2009). URSSA: Evaluating student gains from undergraduate research in science education. *Council on Undergraduate Research Quarterly*, *29*(3), 15–19.

Hurtado, S., Cabrera, N. L., Lin, M. H., Arellano, L., & Espinosa, L. L. (2009). Diversifying Science: Underrepresented Student Experiences in Structured Research Programs. *Research in Higher Education*, *50*(2), 189–214. https://doi.org/10.1007/s11162-008-9114-7

Liberatore, M. J., & Wagner, W. P. (2022). Gender, Performance, and Self-Efficacy: A Quasi-Experimental Field Study. *Journal of Computer Information Systems*, *62*(1), 109–117. https://doi.org/10.1080/08874417.2020.1717397

Linnenbrink-Garcia, L., Perez, T., Barger, M. M., Wormington, S. V., Godin, E., Snyder, K. E., Robinson, K., Sarkar, A., Richman, L. S., & Schwartz-Bloom, R. (2018). Repairing the Leaky Pipeline: A Motivationally Supportive Intervention to Enhance Persistence in Undergraduate Science Pathways. *Contemporary Educational Psychology*, *53*, 181–195. https://doi.org/10.1016/j.cedpsych.2018.03.001

Lopatto, D. (2008). *Classroom Undergraduate Research Experiences Survey*. https://www.grinnell.edu/academics/areas/psychology/assessments/cure-survey

Maton, K. I., & Hrabowski, F. A. (2004). Increasing the number of African American PhDs in the sciences and engineering: A strengths-based approach. *American Psychologist*, *59*, 547–556. https://pubmed.ncbi.nlm.nih.gov/15367090/

U.S. National Science Foundation. (2021). *The STEM Labor Force of Today: Scientists, Engineers, and Skilled Technical Workers*. U.S. National Science Board: Science and Engineering Indicators. https://ncses.nsf.gov/pubs/nsb20212

Nazzaro, V., Rose, J., & Dierker, L. C. (2020). A Comparison of Future Course Enrollment among Students Completing One of Four Different Introductory Statistics Courses. *Statistics Education Research Journal*, *19*(3). https://par.nsf.gov/biblio/10275720.

Pasquale, M. (2015). *Productive Struggle in Mathematics* (Research Brief 2). Education Development Center.

Rowland, S. L., Lawrie, G. A., Behrendorff, J. B., & Gillam, E. M. (2012). Is the undergraduate research experience (URE) always best?: The power of choice in a bifurcated practical stream for a large introductory biochemistry class. *Biochemistry and Molecular Biology Education*, *40*(1), 46–62.

Seymour, E., Hunter, A.-B., Laursen, S. L., & DeAntoni, T. (2004). Establishing the benefits of research experiences for undergraduates in the sciences: First findings from a three year study. *Science Education*, *88*, 493–534.

Weston, T. J., & Laursen, S. L. (2015). The Undergraduate Research Student Self-Assessment (URSSA): Validation for Use in Program Evaluation. *CBE Life Sciences Education*, *14*(3). https://doi.org/10.1187/cbe.14-11-0206

Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology*, *25*(1), 68–81. https://doi.org/10.1006/ceps.1999.1015

Wise, S. L. (1985). The development and validation of a scale measuring attitudes toward statistics. *Educational and Psychological Measurement*, *45*, 401–405.