A Course-Based Undergraduate Research Experience (CURE) in Statistics: Enhancing Educational Outcomes and Promoting Access to Data-Driven Careers

Lisa Dierker, Jennifer Rose, Valerie Nazzaro, Robert Kabacoff, Maryam Gooyabadi, and Manolis Kaparakis
Wesleyan University

Course-based undergraduate research experiences (CUREs) have been shown to help students gain research skills. Few studies, however, have evaluated the impact of CUREs on long-term educational and employment outcomes. This article examines postgraduation outcomes for students enrolling in introductory statistics delivered as a CURE compared with those taking a traditional mathematics-based introductory statistics course. The CURE was found to be associated with a higher likelihood of holding a job in which a primary responsibility was working with data, greater confidence in working with data, and a higher likelihood of earning more than $100K annually. This study suggests that CUREs may not only impart technical skills but also boost students' self-efficacy in applying statistical methods and analyzing data.
Keywords: Course-based undergraduate research experience; STEM; Workforce; Statistics; Curriculum

Most of the top 25 highest paying, in-demand jobs (e.g., data scientist, software architect, analytics manager, IT manager, UX designer, database administrator, solutions architect, etc.) require data skills (Glassdoor, 2016). These and other technology-oriented careers are predicted to rise 13% by 2030 (U.S. Bureau of Labor Statistics, 2022), and the relevance of data skills in many other occupations that have previously been less technology-oriented has also been expanding. Between 2003 and 2013, digitally intensive jobs grew 2.5 times more rapidly than those that do not require digital skills (Lockard & Wolf, 2012). More recent projections include more than half a million job openings in the information technology sector, most from nontechnology-oriented industries that currently employ two thirds of private-sector technology workers (e.g., in manufacturing, retail, and healthcare; U.S. Bureau of Labor Statistics, 2022).

The most promising curricular model believed to promote the kind of data literacy that translates into the job skills needed within the modern analytic workforce is authentic undergraduate research experiences (National Academies of Science, Engineering, and Medicine, 2017). Undergraduate research experiences contribute to the improvement of students' scientific identity, self-efficacy, and content knowledge (Linn et al., 2015). Students who have conducted research as an undergraduate also report gains in scientific problem solving and tolerance for obstacles, findings that are consistent across research field, gender, ethnicity, and type of educational environment (Ahmad &Al-Thani, 2022). These benefits have also been shown to influence career trajectories. In several studies, undergraduate research experiences were associated with an increased interest in STEM careers, increased graduation rates, and the pursuit of advanced degrees (Chamely-Wiik et al., 2023; Haeger et al., 2020; Thiem et al., 2023; Zhuchkova & Bekova, 2023). Despite these consistently positive findings, limitations on time, funding, and space pose constraints, resulting in insufficient research opportunities for undergraduate students (Adebisi, 2022). Additionally, students often lack awareness of the availability of these opportunities, knowledge on how to secure one, or the necessary time outside of regular classes to participate (Bangera & Brownell, 2014).

Course-based undergraduate research experiences (CUREs) avoid many of these limitations by engaging students in research as part of standard curriculum rather than as stand-alone experiences outside of normal coursework (Dolan & Weaver, 2021). In this way, CUREs provide a more inclusive entry point to scientific research by being delivered to larger numbers of students earlier in their college education (Bangera & Brownell, 2014; Watts & Rodriguez, 2023). Because CUREs are embedded within courses, they are more accessible than research internships to students with time-consuming extracurricular activities; to nontraditional students with job and family obligations; to students who attend colleges and universities that are not research-intensive; and to first-generation students who may be less aware of research internships' long-term career value and have fewer soft skills for navigating application pathways for extracurricular research experiences. Instead of exclusively serving advanced STEM students, CUREs are often delivered at the introductory level, targeting both prospective majors and nonmajors (Kortz & van der Hoeven Kraft, 2016). CUREs can therefore exert a greater influence on students' academic and career paths than research experiences that occur later in undergraduate academic programs, after many students have exited the major or withdrawn from college (Auchincloss et al., 2014). An additional advantage of CUREs is that they do not require students to preselect themselves as interested in mathematics or talented in science; CUREs are integrated into existing courses, making them accessible to the widest range of students.

CUREs encompass five key elements, as outlined by Auchincloss et al. (2014):

1. Use of scientific practices: Students engage in activities that mimic the work of professional researchers.
2. Discovery: Students address fresh, unexplored questions, potentially contributing to the field, with both students and instructors unaware of the results.
3. Significance and relevance: Student work adds to the current body of knowledge in the discipline.
4. Collaboration: Students collaborate to solve problems collectively.
5. Iteration: Student work involves iterative scientific practices, such as repeating experiments, revising hypotheses, or devising new approaches to validate findings.

Studies comparing CURE instruction with research internships and laboratory learning experiences have generally found that students report many of the same gains (Olivares-Donoso & González, 2019), including increased scientific identity (Cooper et al., 2020; LaForge & Martin, 2022) and self-efficacy (Cooper et al., 2020; Newell & Ulrich, 2022; Robnett et al., 2015; Teter, 2023; Von der Mehden et al., 2023), improved attitudes toward science (Werth et al., 2023), increased understanding of the scientific method (Flaherty et al., 2017), and greater content knowledge (Flaherty et al., 2017; Gao & Guo, 2023). CUREs help students gain research skills, engage in productive and enjoyable teamwork experiences, and feel motivated and interested in experimental research (Werth et al., 2022).

CUREs have also been shown to promote heightened interest and motivation in science and mathematics as well as improved communication skills (Kortz & van der Hoeven Kraft, 2016). They attract and retain students from underrepresented and marginalized groups (Bangera & Brownell, 2014; Ing et al., 2021), and enhance the retention rates of STEM majors (Hanauer et al., 2017; National Academies of Science, Engineering, and Medicine, 2015; Newell & Ulrich, 2022; Rodenbusch et al., 2016). In the context of statistics education, CUREs have been shown to encourage higher level development of statistical concepts (Moreira da Silva & Pinto, 2014) and to inspire students to engage in statistical inquiry beyond the course (Nazzaro et al., 2020). To date, however, we know of no long-term follow-up studies evaluating the educational and employment outcomes of teaching introductory statistics through a CURE.

Our previous publications have described the development of Passion-Driven Statistics, a multidisciplinary CURE introducing students to the research process through data-driven projects (Dierker et al., 2012). Passion-Driven Statistics is funded by the U.S. National Science Foundation and closely follows the recommendations outlined in the *Course-Based Undergraduate Research Experiences Network* report (Auchincloss et al., 2014) and the *Guidelines for Assessment and Instruction in Statistics Education (GAISE) College Report 2016* (Carver et al., 2016). This research-project-based curriculum teaches students to deploy statistical skills within authentic research projects of students' own choosing, offering individualized hands-on experience in quantitative research and applied statistics through engagement with real-world data and statistical software. The curriculum has been implemented successfully 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 and universities, medical schools, community colleges, and high schools (Dierker, Evia et al., 2018).

The Passion-Driven Statistics CURE has been found to improve enrollment and persistence of historically marginalized groups in STEM. Previous research demonstrated that the CURE enrolls higher rates of historically marginalized students than a mathematics-based statistics curriculum (Dierker et al., 2015). Increased interest in research has also been shown to be higher following enrollment in the project-based course among historically marginalized students compared with nonmarginalized students (Dierker et al., 2016). Students completing the Passion-Driven Statistics CURE are also significantly more likely to report an increase in confidence in managing data, writing syntax or code, and choosing the correct statistical test. CURE students were also more likely than students in a mathematics-based statistics course to show increased interest in pursuing additional opportunities in research and applied statistics (Dierker, Flaming et al., 2018). The Passion-Driven Statistics CURE has also been shown to have an impact on students' academic trajectory. Using causal inference techniques to achieve matched comparisons across three different statistics courses, we determined that students who had 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, or use of statistical software compared with demographically matched students taking either a psychology statistics course or mathematical statistics course (Nazzaro et al., 2020).

The purpose of the present article is to evaluate the long-term impact of the Passion-Driven Statistics CURE by comparing post-undergraduate (post-UG) educational and employment outcomes of individuals completing the CURE as undergraduates compared with those taking a traditional introductory statistics course.

Methods

The CURE

The Passion-Driven Statistics CURE provides an opportunity to use scientific practices by becoming fluent in the work of quantitative research. Allowing flexibility for the discretion of each instructor, this CURE offers the structure for students to (a) read data dictionaries, focusing on topics or constructs that they find relevant and interesting; (b) develop a hypothesis on the basis of the available data; (c) conduct a review of the literature, evaluating previous work related to the hypothesis and demonstrating how the research will contribute to the existing literature; (d) use statistical software (e.g., SAS®, RStudio®, Stata®, Python®, and SPSS®) to examine frequency tables, variability, and central tendency for the selected variables; (e) perform data management on the chosen variables (e.g., setting aside missing data, creating a new variable by collapsing response categories, creating a new variable by aggregating across more than one variable, or labeling variables); (f) graph individual variables and the associations between pairs of variables; (g) evaluate associations with inferential statistical tools; (h) test for moderation (i.e., differences in observed associations with respect to a third variable); (i) build and evaluate multivariate models; and (j) interpret and present results.

The curriculum covers basic themes—such as measuring and interpreting statistical results, choosing analytic techniques that are appropriate for the data, and using software and coding to generate statistics—without neglecting the underlying theory and mathematics. For example, in hypothesis testing, students learn fundamental concepts related to making inferences about populations from samples, including the role of probability; sampling variability and distributions; central limit theorem; significance levels (p-value); and Type I and Type II errors.

To test bivariate hypotheses, students learn to choose analysis of variance (ANOVA), a chi-square test of independence, or Pearson correlation as suitable for their research questions and variable measurement. The ANOVA lesson introduces students to the calculation of *F* as the ratio of the variation among sample means to the variation within groups; the use of boxplots and scatterplots to visualize variation; and interpreting *F* with the associated *p*-value. When making multiple group comparisons, students are taught post-hoc tests to avoid excessive Type I error with an explanation of family-wise Type I error rates and a formula to calculate an adjusted *p* value:

Family-wise error rate = $1 - (1 - \alpha)^n$

where:

α: The significance level for a single hypothesis test

n: The total number of tests

Then, they are shown how to run Tukey's honestly significant difference test and how to interpret the results of differences in means between groups with confidence intervals and associated *p*-values. In addition, great care is taken to present translations of terminology and vocabulary that are used across different disciplines for similar statistical concepts (e.g., independent and dependent variables vs. predictors and outcomes vs. explanatory and response variables).

Passion-Driven Statistics focuses on discovery by taking advantage of students' natural curiosity and providing a common language for approaching questions across numerous disciplines. The CURE is based on original questions and hypotheses that are driven by students' interests. Students begin by examining variables available for the class (from public data sets and codebooks). Students work individually or in pairs to pose a research question that interests them. Some examples include: Is exposure to a drug use prevention curriculum associated with lower rates of experimentation with substances? Is there a relationship between migraine headaches and educational attainment? What factors predict safe-sex practices?

After choosing questions, students conduct a literature review to discover what is already known about their topic of interest. For example, a student might pose the research question "Is child abuse associated with an increased incidence of mental health issues?" and discover that much has been published on this topic. Students are guided in approaches to refine the research question, focus on a subset of the sample, or explore different measurement approaches to allow for novel discoveries. Many projects combine some replication of prior studies with an extension of what is already known. Not uncommonly, students find results that are unexpected or the opposite of what they hypothesized. As an example, a student investigated the role of religiosity in feelings of support during difficult times. The analysis revealed that people who reported strong reliance on their faith and regular attendance of faith-based services also reported the lowest amount of support during difficult times. This was not what the student expected to find and led them to do additional review of the literature to help explain these results. In general, the course encourages students to think critically about their research questions throughout the steps of the research process.

The CURE is also designed to support large numbers of students to learn statistical software and the process of executing a research project through a collaborative model. When possible, and commonly in classes with an enrollment greater than 15–20 students, peers who previously completed the CURE are involved in the mentoring process. Peer mentors may serve as volunteers because of a desire to learn the specific software being taught, be paid through student work programs or training grants, or receive course credit as teaching assistants or through other course designations (e.g., a statistics education practicum). Although each project reflects the individual student's interests and efforts, it is also the result of a collaborative process among students, peer mentors, and the instructor of thinking through each step of the research process. As students engage in productive struggle in the context of doing original research, the instructor and peer mentors support each student individually through ample one-on-one mentoring.

In the classroom, multiple levels of collaboration occur simultaneously. On the student-to-student level, students may learn which of their classmates are working on similar research questions so that those students have the chance to bounce ideas off one another (e.g., decisions about recoding variables). Because the peer mentors have been through the CURE themselves as students, they are also a critical component of the learning environment. They model the thought processes and demonstrate that research can be a trial-and-error process (e.g., through being perplexed by an issue or making coding mistakes and fixing them). Collaboration also serves to help students avoid pitfalls and not flounder on a challenging task. By supporting students as they make decisions about the most appropriate ways to visualize, explore, and analyze data, the CURE emphasizes conceptual understanding and application in the context of authentic research. Rather than focus on rules associated with traditional lists of tools, this CURE proceeds with the decisions and skills involved in quantitative research.

The Passion-Driven Statistics CURE also lays a foundation for understanding the iterative nature of science, beginning with an introduction to data sets and conducting a literature review of primary source research articles. These first components of the CURE teach students to understand the cumulative processes through which data are collected and research studies are conducted, while prompting them to use this insight to develop their own research questions using data codebooks and literature review. Subsequently, data management, exploratory data analysis, and inferential analysis enable preliminary hypothesis testing and implicate additional variables and measurement approaches. During data management and exploratory data analysis, variable measurement is frequently revised to distinguish the research from existing studies identified in the literature review. Students may also identify empirical reasons (e.g., small cell sizes) to revise construct measurement as a result of their data analysis. Further, the Passion-Driven Statistics CURE teaches moderation and confounding to help students incorporate additional variables, rule out alternative explanations, and increase confidence in the findings. As students hone variable measurement and build models, each iteration may require revisiting the literature to ensure that their approach is novel, additional data management to recode variables, revised exploratory data analysis and graphing to generate descriptive statistics, and alternative inferential statistical analysis depending on the type of variable being examined.

The concluding portion of the Passion-Driven Statistics CURE engages students in the processes of evaluating their hypotheses to make claims that are supported by evidence. In developing papers or presentations of their research project, students reflect on the numerous iterations and summarize the insights gained using statistics to address research questions on their chosen topics. As with all dissemination of scientific research, students are faced with the final challenge of sharing relevant details and concluding with implications for future research.

Because the Passion-Driven Statistics curriculum provides the opportunity for students to build on current scientific knowledge, it encourages students to share broadly relevant and important discoveries with their local and academic communities. Students have presented their research posters at local, state, national, and international conferences. Commonly, Passion-

Driven Statistics sites hold poster presentation sessions that allow for not only a meaningful summative assessment of student learning but also the opportunity for students to situate their results within the context of previous research. At this concluding event, students describe their process of inquiry, including their premises, decisions made along the way, conclusions, and any barriers that they faced to audiences of peers, instructors, and external experts. In many cases, poster sessions are open to the campus and local community.

Participants

For this study, administrative data were used to recruit alumni who had previously enrolled between fall semester 2009 and spring semester 2018 in one of two undergraduate introductory statistics courses at our university: the multidisciplinary CURE or a statistics course offered through the mathematics department. The CURE introductory statistics course was offered through the Quantitative Analysis Center, a collaborative effort of academic and administrative departments that supports quantitative analysis across the curriculum. The course catalog described Applied Data Analysis as

> a project-based course [in which] you will have the opportunity to answer questions that you feel passionately about through independent research based on existing data. You will develop skills in generating testable hypotheses, conducting a literature review, preparing data for analysis, conducting descriptive and inferential statistical analyses, and presenting research findings. The course offers one-on-one support, ample opportunities to work with other students, and training in the skills required to complete a project of your own design. These skills will prepare you to work in many different research labs across the University that collect empirical data. It is also an opportunity to fulfill an important requirement in several different majors (Wesleyan University, 2022–2023, "QAC201").

The course format was listed as laboratory and one of several code-based statistical software platforms was used in each section (i.e., SAS, RStudio, Stata, SPSS). The assessments were the same across all instructors and included exams, research project components, a research paper, and a final poster project and presentation.

The course catalog described the introductory statistics course offered through the mathematics department, Elementary Statistics, as covering the topics of "organizing data, central measures, measures of variation, distributions, sampling, estimation, conditional probability (Bayes' theorem), hypothesis testing, simple regression and correlation, and analysis of variation" (Wesleyan University, 2022–2023, "MATH132"). The course format was listed as lecture and statistical software was not generally used. The assessments varied by instructor but commonly included exams, homework, and applied exercises.

Both courses were open to all students and had no prerequisites for enrollment. Each could be used as one option to fulfill a major requirement for Biology, Neuroscience & Behavior, and Psychology. In addition, each could be applied to the natural sciences and mathematics general education recommendation. In addition, the CURE statistics course was an option for fulfilling a major requirement in Government and Sociology and the mathematics course could be used as a requirement for the Molecular Biology & Biochemistry major. Importantly, neither course was specifically required of any student, nor did either represent a single option for fulfilling requirements for any major within the university. A considerable variability in majors existed among the 284 students who had declared a major at the time of the taking the course, so much so that this variable could not be easily quantified. The total sample contained 36 different majors, with 31 different majors in each type of course. Double and even triple majors were not uncommon, with $n = 117$ (41.1%) reporting having more than one major. However, the distribution of majors was similar in both courses, and included majors in the arts and humanities, natural sciences, mathematics, and social and behavioral sciences. Only three students enrolled in the CURE and three students enrolled in the mathematical statistics course had declared mathematics as their major. Enrollments per class section were about 25 students for the CURE course and about 35 students for the mathematics course.

Materials

Administrative data drawn from students' application to the university and class enrollment each semester were supplied by the institutional research office. These data included age at university enrollment, gender, and international student status. Incoming students self-reported also whether they were the first generation of their family to attend college. Financial aid was measured by support in the form of a Pell grant.[1] Self-reported race and ethnicity included endorsement of one or more non-mutually exclusive categories. Variables were created

[1] Student eligibility for this grant is decided by the Free Application for Federal Student Aid and is based on financial need.

indicating year of enrollment in the undergraduate introductory statistics course (1st, 2nd, 3rd or 4th year of college) and number of years since graduation when completing the survey.

Additional undergraduate courses taken following the introductory statistics course that emphasized statistical concepts, research methods, applied data analysis, and use of statistical software were also examined. These included courses offered by the Quantitative Analysis Center[2]; the mathematics department (i.e., Introduction to Probability, Mathematical Statistics); the economics department (i.e., Econometrics); government (i.e., Empirical Methods for Poli Sci, Political Science by Numbers, Media Analysis Research); psychology (i.e., Quantitative Research Methods in Clinical Psychology); biology (i.e., Quantitative Methods); and physics (i.e., Thermal Statistical Physics). Additional data-oriented coursework was measured by whether the student took any additional undergraduate course (yes/no) and as the number of additional courses taken.

Questions on the alumni survey included whether participants completed a degree or certificate in addition to their undergraduate degree and whether they were pursuing any degrees or certificates as a part-time or full-time student. Respondents were also asked their current or most recent job title and the primary responsibilities of their position. With respect to their current or most recent job, a series of questions was asked about how often they work with or use quantitative data: collecting data, analyzing data, communicating the results of data, making decisions based on data. They were also asked how often they would like to work with data in the future (1 = *never* to 5 = *very frequently*) and how confident they feel about working with data (1 = *not at all* to 5 = *extremely confident*). General questions included how many hours participants currently work per week on average, how satisfied they are with the number of hours they are currently working (1 = *not satisfied* to 5 = *very satisfied*) and their salary range (1 = *$0–$24,999* to 8 = *$200,000 and above*). Finally, participants were asked whether they remember taking the target course (CURE or mathematical statistics; title and course catalog description

[2] Business Modeling with Excel, Data Journalism: Intro, Data Science Primer, Data Visualization I, EDA & Pattern Discovery, Experimental Design & Causal Inference, GIS Proseminar; Intro Data Management, Intro Statistical Consulting, Introduction to Text Mining, Hierarchical Linear Models, Intro Bayesian Analysis, Longitudinal Data Analysis, Network Analysis, Latent Variable Analysis, Machine Learning–Data Mining, Network Analysis; Proseminar: Web Scraping, Spatial Data Analysis & Visualization, Statistics Education Practicum; Individual Tutorial, Undergrad Student Forum, Survival Analysis, Teaching Apprentice Tutorial, Working with Excel and VBA, Working with Mathematica, Working with Python, Working with R, Working with SAS, Working with SQL Databases, and Working with Stata.

were included) and whether they think that the target course was useful in preparing them for a job or for graduate school (rated 1 = *less useful than other college courses*; 2 = *similarly useful compared to other college courses*; 3 = *more useful than other college courses*).

Procedure

The study received exempt approval from the Wesleyan University Institutional Review Board (Project ID: 20190701), thus waiving the requirement of informed consent. Only alumni for whom one of these introductory courses was their first college-level statistics course and who were not concurrently enrolled in any other data analysis course were part of the recruitment sample (n = 1,882). Survey links were sent through university email addresses and LinkedIn messaging between July 2021 and September 2022 (1 to 11 years post-UG). Alumni were given a description of the research and were informed that their participation would be confidential. Individuals could freely choose to participate. Participants gave consent to participate by clicking on the link to the online survey and completing the survey. The survey took less than 5 min to complete, and participants were offered a $25 Amazon gift certificate for completion. Data were identifiable through a student ID number to enable us to merge survey responses with administrative data. All data were then de-identified and given a random unique identifier before analysis. A total of 907 alumni (551 CURE statistics students and 356 mathematical statistics students) completed the survey (Table 1).

Table 1

Demographic and Background Characteristics of Respondents

Variable	CURE statistics n = 551	Mathematical statistics n = 356
Age at university enrollment	M = 18.8 (SD = 1.7)	M = 18.6 (SD = 0.7)
Gender		
Women	312 (56.6%)	215 (60.4%)
Men	239 (43.4%)	141 (39.6%)
Ethnicity		
Hispanic	62 (11.3%)	42 (11.8%)
Black	67 (12.2%)	32 (9.0%)

Native Hawaiian/Pacific Islander	1 (0.2%)	3 (0.8%)
American Indian/Native American	9 (1.6%)	1 (0.3%)
Asian	104 (18.9%)	72 (20.2%)
White	338 (61.3%)	218 (61.2%)
Other	3 (0.5%)	1 (0.3%)
International student	50 (9.1%)	21 (5.9%)
First-generation college student	91 (16.5%)	61 (17.1%)
Pell grant eligible	106 (19.2%)	66 (18.5%)
Year of enrollment in statistics course		
1st year	60 (10.9%)	142 (39.9%)
2nd year	219 (39.8%)	122 (34.3%)
3rd year	155 (28.1%)	52 (14.6%)
4th year	117 (21.2%)	40 (11.2%)
Years since graduation	$M = 5.3$ ($SD = 2.6$)	$M = 4.5$ ($SD = 2.8$)

Because of the quasi-experimental design of our study, we used a causal inference technique to achieve matched comparison groups. A case-control matched analysis is often used in observational studies to reduce selection bias and approximate a randomized trial (Rosenbaum, 2010). A propensity score is the predicted probability of an outcome. Studies have shown that a sample matched on propensity score will be similar for all the covariates included in the calculation of the propensity score. Thus, matching on the propensity score can reduce the selection bias in an observational study (Rosenbaum, 2010). The balance achieved by matching survey respondents on a comprehensive range of potential confounding demographic characteristics is similar to the balance that is achieved through random assignment into treatment groups. Although propensity score matching does not fully eliminate the potential for confounding, it significantly reduces that potential in much the same way as randomization. In the absence of randomization, propensity score matching is an effective method for reducing the potential for confounding even when samples appear homogeneous.

In our analyses, students completing the CURE statistics course were considered the treatment group compared with the control group of students completing the mathematical statistics course. Propensity scores were calculated using student background characteristics (i.e., race, gender, age, first-generation college student status, Pell grant eligibility, international student status) as well as college year of enrollment in the introductory statistics course and number of years since graduation (1 through 10.5). Students were matched on the propensity for enrollment in the CURE statistics course using the OneToManyMTCH SAS Macro (Parsons, 2004), specifying a 1:1 nearest neighbor match. The match resulted in balance on each of the matching variables, that is, significant differences no longer existed between treatment and control conditions on each variable. Using this technique created a data set in which each graduate who had taken the CURE statistics course was matched to a student who had taken the mathematics statistics course. Generating the matched sample successfully eliminated differences on all background variables (i.e., all differences after matching resulted in $p > .05$), resulting in a matched sample of 272 survey respondents having taken each course.

Chi-square tests of independence and ANOVA were then used for categorical and quantitative outcomes, respectively. Path analysis using Mplus (version 8.8) was then conducted on the matched samples using maximum likelihood estimation (and a logit link function for binary outcomes) to evaluate the mediating effect of taking additional data- and research-oriented courses at the undergraduate level on the association between enrollment in the CURE and educational and workforce outcomes. Path analysis is a structural equation modeling technique that quantifies the mediating role of one or more variables in the association between an independent variable and a dependent variable. It estimates the direct associations between independent and dependent variables, as well as the impact of potential mediators on these direct associations. Next, Cohen's *d* effect sizes were calculated for all associations. Finally, odds ratios were transformed to Cohen's *d* estimates by dividing the log odds ratio by 1.81, as outlined by Chinn (2000). The log odd ratios of dichotomous variables were transformed into an effect size (Cohen's *d*) to perform an overall meta-analysis of effect size (Chinn, 2000). The conversion of an odds ratio to a Cohen's *d* effect size estimate was calculated by dividing the odds ratio by 1.81 (Chinn, 2000).

Results

Measured outcomes by matched sample are presented in Table 2. Alumni who had taken the CURE statistics course as an undergraduate were more likely to take at least one additional data-oriented course as well as a larger number of data-oriented courses before completing their bachelor's degree compared with those who had taken the mathematical statistics course. CURE students also reported their current confidence in working with data as higher, were more likely to currently earn more than $100K annually and were more likely to describe their current work as including data collection, analysis, or interpretation as a primary responsibility.

Table 2

Undergraduate (UG) Courses and Post-UG Outcomes for the Matched Samples

Variable	CURE Statistics $n = 272$	Math Statistics $n = 272$	Statistics	Cohen's d effect size
		UG data courses		
Took additional data-oriented courses as UG	148 (54%)	85 (31.3%)	$\chi^2(1) = 29.8$, $p < .001$	0.44
Number of data-oriented courses taken as UG following their introductory statistics course	$M = 2.2$ ($SD = 1.7$)	$M = 1.5$ ($SD = 1.0$)	$F(1, 542) = 36.4$, $p < .001$	0.52
		Post-UG education and work		
Educational Attainment				
Completion of advanced degree/certificate beyond BA/BS	107 (39.3%)	99 (36.4%)	$\chi^2(1) = 0.5$, $p = .479$	0.06
Currently pursuing degree or certificate as PT/FT student	66 (24.3%)	86 (31.6%)	$\chi^2(1) = 3.7$, $p = .056$	0.15
Data as primary job responsibility in current/most recent job	55 (20.2%)	33 (12.1%)	$\chi^2(1) = 6.6$, $p = .010$	0.20
Frequency of working with data in current/most recent job (1 = *never* to 5 = *very frequently*)				
Collecting data	$M = 3.3$ ($SD = 1.3$)	$M = 3.4$ ($SD = 1.3$)	$F(1, 540) = 0.3$, $p = .580$	0.05
Analyzing data	$M = 3.6$	$M = 3.4$	$F(1, 540) = 2.2$,	0.13

	(SD = 1.3)	(SD = 1.4)	p = .141	
Communicating results of data	M = 3.7 (SD = 1.3)	M = 3.7 (SD = 1.3)	$F(1, 541) = 0.2$, p = .660	0.04
Making decisions based on data	M = 3.7 (SD = 1.3)	M = 3.6 (SD = 1.3)	$F(1, 542) = 2.7$, p = .104	0.14
Would like to work with data in the future	M = 3.8 (SD = 1.0)	M = 3.8 (SD = 1.0)	$F(1, 542) = 0.4$, p = .534	0.05
Confidence in working with data (1 = *not at all confident* to 5 = *extremely confident*)	M = 3.5 (SD = 1.0)	M = 3.3 (SD = 1.0)	$F(1, 542) = 4.3$, p = .039	0.18
Average number of hours worked/week	M = 40.4 (SD = 14.7)	M = 37.9 (SD = 16.7)	$F(1, 537) = 3.4$, p = .066	0.16
Satisfaction with number of hours worked/week (1 = *not* to 5 = *very*)	M = 3.4 (SD = 1.2)	M = 3.4 (SD = 1.0)	$F(1, 537) = 0.7$, p = .410	0.07
Salary > $100K/year	94 (34.7%)	69 (25.5%)	$\chi^2(1) = 5.5$, p = .019	0.19
Do you remember taking statistics course	267 (98.2%)	248 (91.5%)	$\chi^2(1) = 12.3$, p < .001	0.3
More useful than other college courses preparing for a job or graduate school	114 (42.1%)	37 (13.8%)	$\chi^2(1) = 53.4$, p < .001	0.6

When asked about their experience with the introductory statistics course, those taking the CURE statistics course were more likely to remember it and to judge it as more useful than other college courses in preparing them for a job or graduate school. No differences were found between the matched samples with respect to having completed a degree or certificate program beyond their undergraduate degree or being enrolled as a full-time or part-time student. Self-reported frequency of working with data in the context of collecting data, analyzing data, communicating the results of data, making decisions based on data, and how often they would like to work with data in the future were also statistically similar between groups. No differences emerged in the average number of hours worked each week or in respondents' satisfaction with the number of hours worked.

To test for mediation, significant bivariate associations must exist between the predictor (enrollment in an introductory statistics course), the mediator (enrollment in additional data-oriented courses), and the post-UG outcomes. These conditions were met for all the significant post-UG outcomes in Table 2, with the exception of remembering the statistics course, which was not significantly associated with enrollment in additional data-oriented courses. The results of the path models evaluating the mediating effect of the number of additional data-oriented courses taken on the association between the CURE statistics course and post-UG outcomes meeting the conditions for mediation are shown in Table 3. The path models revealed that the total effects of the CURE statistics course and taking additional data-oriented courses on post-UG outcomes were all statistically significant and positive. Odd ratios (*OR*s) ranged from 1.5 (95% CI [1.1, 2.2]) for currently earning more than $100k a year ($d = 0.23$) to $OR = 4.5$ (95% CI [3.0, 7.1]) for finding the CURE statistics course more useful than other college courses in preparing for a job or graduate school ($d = 0.83$). Effect sizes indicated that the total effects were small, with the exception of finding the CURE statistics course more useful, for which the results suggested a larger effect.

Table 3

Mediating, Direct, and Total Effects for CURE Statistics Course on Post-UG Outcomes

Post-UG outcome	Direct effect of CURE statistics	Indirect effect of additional courses	Total effect (direct and indirect combined)
Data as primary job responsibility in current/most recent job	$OR = 1.6$, 95% CI [1.0, 2.7], $p = .043$, $d = 0.27$	$OR = 1.1$, 95% CI [1.0, 1.2], $p = .053$, $d = 0.06$	$OR = 1.8$, 95% CI [1.1, 3.0], $p = .012$, $d = 0.33$
Salary > $100K/year	$OR = 1.2$, 95% CI [0.8, 1.9], $p = .294$, $d = 0.11$	$OR = 1.3$, 95% CI [1.1, 1.4], $p < .001$, $d = 0.12$	$OR = 1.5$, 95% CI [1.1, 2.2], $p = .024$, $d = 0.23$
More useful than other college courses preparing for a job or graduate school	$OR = 3.6$, 95% CI [2.4, 5.7], $p < .001$, $d = 0.71$	$OR = 1.2$, 95% CI [1.1, 1.4], $p = .003$, $d = 0.12$	$OR = 4.5$, 95% CI [3.0, 7.1], $p < .001$, $d = 0.83$

Confidence in working with data	B = 0.1, 95% CI [−0.1, 0.22], p = .534, d = 0.10	B = 0.1, 95% CI [0.1, 0.2], $p < .001$, d = 0.24	B = 0.2, 95% CI [0.0, 0.3], p = .034, d = 0.37

Confidence in working with data was also greater for CURE students who took at least one additional data-oriented course (B = 0.2, 95% CI [0.0, 0.3]), although the effect was small (d = 0.20). The independent direct effects of the CURE statistics course after taking into account the mediating effect of taking additional data-oriented course(s) were significant, but small (d = 0.27) for having data as a primary responsibility in the current/most recent job, and moderate for finding the CURE course more useful than other college courses in preparing for a job or graduate school (d = 0.71). Small but statistically significant mediating effects of taking additional data-oriented courses indicated that the effect of taking the CURE statistics course on post-UG outcomes was partially mediated by taking additional statistics or data-oriented courses for earning more than $100K a year (*OR* = 1.3, 95% CI [1.1, 1.4], d = 0.12), finding the CURE course more useful than other college courses in preparing for a job or graduate school (*OR* = 1.2, 95% CI [1.1, 1.4], d = 0.12), and having greater confidence in working with data (B = 0.1, 95% CI [0.1, 0.2], d = 0.24).

Discussion

The present study is the first to evaluate the potential impact of the Passion-Driven Statistics curriculum on post-UG educational and employment outcomes. Findings suggest that enrollment in this course-based undergraduate research experience (CURE) is associated with a higher likelihood of holding a job in which a primary responsibility includes working with data, greater self-reported confidence in working with data, and a higher likelihood of earning more than $100K annually. CURE students were also more likely to remember their introductory statistics course and to judge it as more useful than other college courses in preparing them for a job or graduate school. However, all these associations were small, with the exception of finding the CURE course more useful, which had a moderate effect size.

The mediation findings reveal that the estimated effects of CURE can be partially attributed to the indirect effects of taking additional statistics or data-oriented courses on several of the post-UG outcomes, although these mediating effects were small. Whereas students who enrolled in the CURE are expected to be more likely to make salaries greater than $100K annually and are also more confident in working with data, one mechanism through which enrollment in the CURE potentially affects these outcomes is due to some extent to the CURE influencing students to take additional data courses. Taking additional courses was not found to be necessary for CURE students to hold a job in which a primary responsibility includes working with data or to have found that the course was more useful than other courses in preparing them for a job or graduate school. Given that participation by upperclassmen in introductory statistics courses provides less opportunity to influence students' academic trajectories, more needs to be done to encourage enrollment in research project-based courses as early as possible (Carver et al., 2016) and to offer those experience across a range of general education opportunities and requirements.

Identifying appropriate comparison groups in this type of research can be challenging. A strength of this study is our ability to match a comparison group on diverse demographic variables as well as the year in college that respondents took the introductory course and the number of years since earning their bachelor's degree. As with any nonexperimental design, the possibility exists of lurking variables that could confound observed associations; however, this risk was significantly reduced by matching on a wide range of potentially confounding demographic characteristics. Note that longitudinal research following students postgraduation is much rarer than simple cross-sectional analysis because of the resources required as well as the difficulty in reaching students and obtaining an adequate response rate. Thus, this study provides new evidence that increases confidence that the nature of college course experiences may have long-term effects. To date, studies evaluating longer term outcomes of undergraduate research have generally focused on research internships rather than research experiences embedded within introductory courses (Cooper et al., 2019; Mátrai et al., 2022; Sell et al., 2018) and few have included a comparison group (Thiem et al., 2023).

Many educational settings have created extensively structured requirements that give students little choice or flexibility that might uncover preferences and motivation for course-based research experiences. Though it was quasi-experimental in design, a strength of the present study is that it was conducted at a university with an open curriculum (i.e., no requirements other than completing a major and a total of 32 college level courses) in which students across different majors could pursue a range of data- and research-oriented courses following the introductory statistics course in the service of their interest and motivation. Currently, the Passion-Driven Statistics CURE is being used in numerous educational contexts including liberal arts colleges, large state universities, regional colleges or universities, community colleges, and high schools. Future research focusing on more homogenous groups such as 1st-year students and students pursuing specific majors is needed to evaluate the long-term outcomes of exposure to the Passion-Driven Statistics CURE and to work toward replication of these effects across diverse secondary and postsecondary educational settings (Dierker, Evia et al., 2018) as well as diverse groups of students (Dierker et al., 2015).

Some scholars have claimed that affecting students' attitudes is the second most important outcome for an introductory statistics course, after increasing conceptual understanding (Ramirez et al., 2012). Negative feelings about previous experiences with statistics is associated with a decrease in the likelihood that students will make use of statistical information to support evidence-based practice, take additional coursework, or pursue a career that requires statistical literacy (Cladera et al., 2019). The present findings suggest that a CURE given in the context of an introductory statistics course may foster the interest and motivation needed to pursue additional coursework and, in turn, facilitate entry into the data-analytics economy.

Despite these strengths, the differences found between the introductory statistics experiences may be explained by unmeasured differences between courses (e.g., research-project-based approach, flipped classroom, intensive use of statistical software, etc.). Therefore, our findings do not suggest that any instructional method on its own increases the likelihood of entering the data-analytics economy. Notably, students completing the CURE have been previously shown to be more likely to take additional data- and research-oriented courses compared with a matched sample of students completing an activity-based introductory statistics course that emphasized the use of statistical software (Nazzaro et al., 2020). This provides at least preliminary evidence that the independent research aspect of the CURE, rather than reliance on software alone, encourages enrollment in additional coursework. Still, no study can claim that all potential confounding factors have been erased by the use of propensity score matching. Note, too, that our methodology achieved balance only between the observed covariates. Other potentially confounding variables were not accounted for in this approach, and our attempt to estimate the causal effect of the course may be limited by these unobserved covariates (e.g., high school experience with research or prior interest in statistics and data analysis). Further research will be needed to replicate and extend the present findings.

Statistical analysis plays a significant role across the natural and social sciences, and, more recently, has become important to scholars within the humanities (Gaffield, 2018). Yet, most empirical studies examining the value of undergraduate research experiences have focused on students within STEM fields alone. Given the ubiquitous nature of data across industries as well as trends in college course selection showing that students continue to overwhelmingly pursue undergraduate degrees in psychology, education, business, economics, sociology, and political science (Irwin et al., 2022), the promotion of equity and inclusion will need to involve opportunities that expose large numbers of students pursuing diverse degrees to conducting independent research early in their academic careers. We have previously demonstrated that our statistics CURE attracts higher rates of underrepresented minority (URM) students compared with a traditional introductory statistics curriculum (Dierker et al., 2015), and although URM students have considered the material covered in the CURE more difficult than non-URM students, they have demonstrated similar levels of increased confidence in applied skills and interest in follow-up courses. CUREs can be a vehicle to improve URM student outcomes, but more research is needed to understand their impacts on student populations that are diverse with respect not only to race but also to educational background.

Additionally, research examining dose response is critically needed. How many CURE experiences over what duration produce the most positive learning and employment outcomes (DeChenne-Peters et al., 2023)? Designing a single course that would prepare students for the amount or complexity of data they will encounter in their careers and as citizens would be impossible (Horton, 2015), but courses that lay the foundation for such exploration are of vital importance. Curricula need to be designed in a way that imparts a deep interest among students and a desire to continue learning. Creating an introductory statistics course that leaves students eager to continue to engage in statistical thinking may be the most important goal. Because CUREs can be implemented in normally scheduled classes, they lower the barriers inherent in traditional undergraduate research experiences and represent an important vehicle for improving undergraduate outcomes and preparing students for the modern data-analytic workforce (Bhattacharyya et al., 2020). The present findings strongly suggest that courses that promote further study beyond the introductory level will help solidify and broaden data literacy and that a course in which students answer questions of personal interest through independent research fosters the necessary interest and motivation. College courses that are modified to include authentic research experiences may have an impact on both a student's educational trajectory as well as their future employment opportunities.

Most importantly, these results add to the accumulating evidence that students benefit from research experiences and that these experiences can be provided within introductory courses. More work is required to further explore the impact of CUREs on increasing graduation rates, reducing years to graduation, improving discipline-specific retention, and broadening inclusion of students from diverse backgrounds (Bell et al., 2017). Currently, the opportunity to conduct original research in mathematics courses, particularly at the introductory level, is rare. Though mathematics faculty have embraced a number of engaged learning practices at the undergraduate level (Braun et al., 2017), opportunities to involve students in authentic research experiences remain largely elusive. For those students who are able to access authentic research opportunities, these often occur late in the college experience, limiting their impact on educational development.

We suggest that a first step in integrating research opportunities into introductory mathematics courses should include efforts to align the curriculum with current industry trends, including mathematical skills and competencies that are in demand in modern analytics industries. This can be done by reviewing job postings and descriptions, consulting with employers and industry experts, and surveying alumni. Armed with this information, the possibility will emerge to reimagine and redesign formative and summative assessments by aligning them to standards and benchmarks that mirror current industry trends. Collecting this data enables educational institutions to tailor curricula to match the evolving needs of our students, thereby bridging the gap between academic learning and practical application.

It will also be important for future research to identify essential aspects of CUREs and overcome barriers to their implementation. Whereas some individual instructors or institutions excel in CUREs, a lack of large-scale implementation remains. Additional work is needed to better understand the features of CUREs that can be best integrated into courses across the mathematics curriculum. Still, mathematics can play a vital role in helping to overcome adoption barriers and advocate for the development of a comprehensive resource to guide the design, implementation, and assessment of CUREs. Recognizing the significant benefits to students, we stress the importance of the mathematics and statistics community in fostering sustained research experiences. We urge the *JRME* readership to expand CURE offerings and to make research a standard practice within their introductory curriculum, engage in evaluation and process improvement and explore the factors contributing to their effectiveness in preparing students for the modern data-analytics landscape.[3]

Conclusion

The present article contributes to current knowledge in several ways. First, an introductory statistics CURE is associated with the development of skills and knowledge relevant to data-driven careers and enhances students' professional opportunities. Second, our findings suggest that CUREs may not only impart technical skills but also boost students' self-efficacy in applying statistical methods and analyzing data. Finally, the present results highlight the potential economic benefits associated with engaging in authentic research experiences early in one's undergraduate education. Overall, the findings contribute to a growing body of evidence supporting the effectiveness of CUREs in promoting access to STEM careers and enhancing students' educational and employment outcomes.

References

Adebisi, Y. A. (2022). Undergraduate students' involvement in research: Values, benefits, barriers and recommendations. *Annals of Medicine & Surgery*, *81*, Article 104384. https://doi.org/10.1016/j.amsu.2022.104384

Ahmad, Z., & Al-Thani, N. J. (2022). Undergraduate research experience models: A systematic review of the literature from 2011 to 2021. *International Journal of Educational Research*, *114*, Article 101996. https://doi.org/10.1016/j.ijer.2022.101996

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

[3] The dissemination of the Passion-Driven Statistics CURE has proceeded through the sharing of open educational resources, and supporting materials can assist in the analysis phase of this as well as other CUREs in which data is used. Resources are available at https://passiondrivenstatistics.com. See also our free e-book at (http://bit.ly/Passion-DrivenStatistics), with links to videos that allow instructors to flip the classroom and translation code (http://bit.ly/PDS_TranslationCode) aimed at supporting the use of statistical software across all the major platforms.

Bangera, G., & Brownell, S. E. (2014). Course-based undergraduate research experiences can make scientific research more inclusive. *CBE—Life Sciences Education*, *13*(4), 602–606. https://doi.org/10.1187/cbe.14-06-0099Bell, J. K., Eckdahl, T. T., Hecht, D. A., Killion, P. J., Latzer, J., Mans, T. L., ... & Ellis Bell, J. (2017). CUREs in biochemistry—where we are and where we should go. *Biochemistry and Molecular Biology Education*, 45(1), 7–12. doi: 10.1002/bmb.20989

Bhattacharyya, P., Chan, C. W. M., Duchesne, R. R., Ghosh, A., Girard, S. N., & Ralston, J. J. (2020). Course-based research: A vehicle for broadening access to undergraduate research in the twenty-first century. *Scholarship and Practice of Undergraduate Research*, *3*(3), 14–27. https://doi.org/10.18833/spur/3/3/7

Braun, B., Bremser, P., Duval, A. M., Lockwood, E., & White, D. (2017). What does active learning mean for mathematicians? *Notices of the American Mathematical Society*, *64*(2), 124–129. https://doi.org/10.1090/noti1472

Carver, R., Everson, M., Gabrosek, J., Horton, N., Lock, R., Mocko, M., Rossman, A., Holmes Rowell, G., Velleman, P., Witmer, J., & Wood, B. (2016). *Guidelines for assessment and instruction in statistics education (GAISE) college report 2016*. American Statistical Association. https://www.amstat.org/docs/default-source/amstat-documents/gaisecollege_full.pdf

Chamely-Wiik, D., Ambrosio, A., Baker, T., Ghannes, A., & Soberon, J. (2023). The impact of undergraduate research experience intensity on measures of student success. *Journal of the Scholarship of Teaching and Learning*, *23*(1), 14–30. https://doi.org/10.14434/josotl.v23i1.32675

Chinn, S. (2000). A simple method for converting an odds ratio to effect size for use in meta-analysis. *Statistics in Medicine*, *19*(22), 3127–3131. https://doi.org/10.1002/1097-0258(20001130)19:22<3127::AID-SIM784>3.0.CO;2-M

Cladera, M., Rejón-Guardia, F., Vich-i-Martorell, G. À., & Juaneda, C. (2019). Tourism students' attitudes toward statistics. *Journal of Hospitality, Leisure, Sport & Tourism Education*, *24*, 202–210. https://doi.org/10.1016/j.jhlste.2019.03.002

Cooper, K. M., Gin, L. E., Akeeh, B., Clark, C. E., Hunter, J. S., Roderick, T. B., Elliott, D. B., Gutierrez, L. A., Mello, R. M., Pfeiffer, L. D., Scott, R. A., Arellano, D., Ramirez, D., Valdez, E. M., Vargas, C., Velarde, K., Zheng, Y., & Brownell, S. E. (2019). Factors that predict life sciences student persistence in undergraduate research experiences. *PLOS ONE*, *14*(8), Article e0220186. https://doi.org/10.1371/journal.pone.0220186

Cooper, K. M., Knope, M. L., Munstermann, M. J., & Brownell, S. E. (2020). Students who analyze their own data in a course-based undergraduate research experience (CURE) show gains in scientific identity and emotional ownership of research. *Journal of Microbiology & Biology Education*, *21*(3). https://doi.org/10.1128/jmbe.v21i3.2157

DeChenne-Peters, S. E., Rakus, J. F., Parente, A. D., Mans, T. L., Eddy, R., Galport, N., Koletar, C., Provost, J. J., Bell, J. E., & Bell, J. K. (2023). Length of course-based undergraduate research experiences (CURE) impacts student learning and attitudinal outcomes: A study of the Malate dehydrogenase CUREs Community (MCC). *PLOS ONE*, *18*(3), Article e0282170. https://doi.org/10.1371/journal.pone.0282170

Dierker, L., 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), Article 2. https://doi.org/10.20429/ijsotl.2016.100102

Dierker, L., Cooper, J., 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., Evia, J. R., Singer-Freeman, K., Woods, K., Zupkus, J., Arnholt, A., Moliski, E. G., Deckard, N. D., Gallagher, K., & Rose, J. (2018). 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/10.20448/2003.32.52.64

Dierker, L., Flaming, K., Cooper, J. K., Singer-Freeman, K., Germano, K., & Rose, J. (2018). 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*(1), 16–28. https://doi.org/10.33094/6.2017.2018.21.16.28

Dierker, L., 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.

Dolan, E. L., & Weaver, G. C. (2021). *A guide to course-based undergraduate research: Developing and implementing CUREs in the natural sciences*. W. H. Freeman & Macmillan Learning.

Flaherty, E. A., Walker, S. M., Forrester, J. H., & Ben-David, M. (2017). Effects of course-based undergraduate research experiences (CURE) on wildlife students. *Wildlife Society Bulletin*, *41*(4), 701–711. https://doi.org/10.1002/wsb.810

Gaffield, C. (2018). Words, words, words: How the digital humanities are integrating diverse research fields to study people. *Annual Review of Statistics and Its Application*, *5*, 119–139. https://doi.org/10.1146/annurev-statistics-031017-100547

Gao, L., & Guo, M. (2023). A course-based undergraduate research experience for bioinformatics education in undergraduate students. *Biochemistry and Molecular Biology Education*, *51*(2), 189–199. https://doi.org/10.1002/bmb.21710

Glassdoor. (2016, November 8). 25 highest paying jobs in demand. *Glassdoor Blog*. https://www.glassdoor.com/blog/highest-paying-jobs-demand

Haeger, H., Banks, J. E., Smith, C., & Armstrong-Land, M. (2020). What we know and what we need to know about undergraduate research. *Scholarship and Practice of Undergraduate Research*, *3*(4), 62–69. https://doi.org/10.18833/spur/3/4/4

Hanauer, D. I., Graham, M. J., SEA-PHAGES, Betancur, L., Bobrownicki, A., Cresawn, S. G., Garlena, R. A., Jacobs-Sera, D., Kaufmann, N., Pope, W. H., Russell, D. A., Jacobs, W. R., Jr., Sivanathan, V., Asai, D. J., & Hatfull, G. F., (2017). An inclusive research education community (iREC): Impact of the SEA-PHAGES program on research outcomes and student learning. *Proceedings of the National Academy of Sciences*, *114*(51), 13531–13536. https://doi.org/10.1073/pnas.1718188115

Horton, N. J. (2015). Challenges and opportunities for statistics and statistical education: Looking back, looking forward. *The American Statistician*, *69*(2), 138–145. https://doi.org/10.1080/00031305.2015.1032435

Ing, M., Burnette, J. M., III, Azzam, T., & Wessler, S. R. (2021). Participation in a course-based undergraduate research experience results in higher grades in the companion lecture course. *Educational Researcher*, *50*(4), 205–214. https://doi.org/10.3102/0013189X20968097

Irwin, V., De La Rosa, J., Wang, K., Hein, S., Zhang, J., Burr, R., Roberts, A., Barmer, A., Bullock Mann, F., Dilig, R., & Parker, S. (2022). *The condition of education 2022* (Report No. NCES 2022-144). U.S. Department of Education, National Center for Education Statistics. https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2022144

Kortz, K. M., & van der Hoeven Kraft, K. J. (2016). Geoscience education research project: Student benefits and effective design of a course-based undergraduate research experience. *Journal of Geoscience Education*, *64*(1), 24–36. https://doi.org/10.5408/15-11.1

LaForge, J., & Martin, E. C. (2022). Impact of authentic course-based undergraduate research experiences (CUREs) on student understanding in introductory biology laboratory courses. *The American Biology Teacher*, *84*(3), 137–142. https://doi.org/10.1525/abt.2022.84.3.137

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

Lockard, C. B., & Wolf, M. (2012, January). Employment outlook 2010–2020: Occupational employment projections to 2020. *Monthly Labor Review*, 84–108. https://www.bls.gov/opub/mlr/2012/01/art5full.pdf

Mátrai, E., Kwok, S. T., Grainger, D., Boos, M., & Pogány, Á. (2022). The academic, career and personal benefits of zoological internships: A 10-year evaluation of a dolphin research internship. *Journal of Zoo and Aquarium Research*, *10*(1), 31–39. https://doi.org/10.19227/jzar.v10i1.660

Moreira da Silva, M. P., & Pinto, S. S. (2014). Teaching statistics through learning projects. *Statistics Education Research Journal*, *13*(2), 177–186. https://doi.org/10.52041/serj.v13i2.289

National Academies of Science, Engineering, and Medicine. (2015). *Integrating discovery-based research into the undergraduate curriculum: Report of a convocation*. National Academies Press. https://doi.org/10.17226/21851

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

Nazzaro, V., Rose, J., & Dierker, L. (2020). A comparison of future course enrollment among students completing one of four different introductory statistics courses. *Statistics Education Research Journal*, *19*(3), 6–17. https://doi.org/10.52041/serj.v19i3.53

Newell, M. J., & Ulrich, P. N. (2022). Gains in scientific identity, scientific self-efficacy, and career intent distinguish upper-level CUREs from traditional experiences in the classroom. *Journal of Microbiology & Biology Education*, *23*(3), Article e00051-22. https://doi.org/10.1128/jmbe.00051-22

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

Parsons, L. S. (2004). Performing a 1:N case-control match on propensity score. In D. Owen (Ed.), *Proceedings of the 29th annual SAS Users Group International Conference* (pp. 165–129). SAS Users Group. https://support.sas.com/resources/papers/proceedings/proceedings/sugi29/165-29.pdf

Ramirez, C., Schau, C., & Emmioğlu, E. (2012). The importance of attitudes in statistics education. *Statistics Education Research Journal*, *11*(2), 57–71. https://doi.org/10.52041/serj.v11i2.329

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

Rodenbusch, S. E., Hernandez, P. R., Simmons, S. L., & Dolan, E. L. (2016). Early engagement in course-based research increases graduation rates and completion of science, engineering, and mathematics degrees. *CBE—Life Sciences Education*, *15*(2), Article 20. https://doi.org/10.1187/cbe.16-03-0117

Rosenbaum, P. R. (2010). *Design of observational studies*. Springer. https://doi.org/10.1007/978-1-4419-1213-8

Sell, A. J., Naginey, A., & Stanton, C. A. (2018). The impact of undergraduate research on academic success. *Scholarship and Practice of Undergraduate Research*, *1*(3), 19–29. https://doi.org/10.18833/spur/1/3/8

Teter, K. (2023). Science communication from a course-based undergraduate research experience. *Biochemistry and Molecular Biology Education*, *51*(4), 370–376. https://doi.org/10.1002/bmb.21728

Thiem, J., Preetz, R., & Haberstroh, S. (2023). How research-based learning affects students' self-rated research competences: Evidence from a longitudinal study across disciplines. *Studies in Higher Education*, *48*(7), 1039–1051. https://doi.org/10.1080/03075079.2023.2181326

U.S. Bureau of Labor Statistics. (2022). *Occupational outlook handbook: Computer and information technology occupations*. U.S. Department of Labor. https://www.bls.gov/ooh/computer-and-information-technology/home.htm

Von der Mehden, B. M., Pennino, E. M., Fajardo, H. L., Ishikawa, C., & McDonald, K. K. (2023). Building authentic science experiences: Students' perceptions of sequential course-based undergraduate research. *CBE—Life Sciences Education*, *22*(4), Article 46. https://doi.org/10.1187/cbe.23-03-0042

Watts, F. M., & Rodriguez, J.-M. G. (2023). A review of course-based undergraduate research experiences in chemistry. *Journal of Chemical Education*, *100*(9), 3261–3275. https://doi.org/10.1021/acs.jchemed.3c00570

Werth, A., West, C. G., & Lewandowski, H. J. (2022). Impacts on student learning, confidence, and affect in a remote, large-enrollment, course-based undergraduate research experience in physics. *Physical Review Physics Education Research*, *18*(1), Article 010129. https://doi.org/10.1103/PhysRevPhysEducRes.18.010129

Werth, A., West, C. G., Sulaiman, N., & Lewandowski, H. J. (2023). Enhancing students' views of experimental physics through a course-based undergraduate research experience. *Physical Review Physics Education Research*, *19*(2), Article 020151. https://doi.org/10.1103/PhysRevPhysEducRes.19.020151

Wesleyan University. (2022–2023). *Course Catalog*. https://owaprod-pub.wesleyan.edu/reg/!wesmaps_page.html?stuid=&term=1229

Zhuchkova, S., & Bekova, S. (2023). Building a strong foundation: How pre-doctorate experience shapes doctoral student outcomes. *PLOS ONE*, *18*(9), Article e0291448. https://doi.org/10.1371/journal.pone.0291448

Authors

Lisa Dierker, Quantitative Analysis Center, Wesleyan University, Middletown, CT 06459; ldierker@wesleyan.edu

Jennifer Rose, Quantitative Analysis Center, Wesleyan University, Middletown, CT 06459; jrose01@wesleyan.edu

Valerie Nazzaro, Quantitative Analysis Center, Wesleyan University, Middletown, CT 06459; vnazzaro@wesleyan.edu

Robert Kabacoff, Quantitative Analysis Center, Wesleyan University, Middletown, CT 06459; rkabacoff@wesleyan.edu

Maryam Gooyabadi, Quantitative Analysis Center, Wesleyan University, Middletown, CT 06459; mGooyabadi@wesleyan.edu

Manolis Kaparakis, Quantitative Analysis Center, Wesleyan University, Middletown, CT 06459; mkaparakis@wesleyan.edu

Submitted June 4, 2023

Accepted pending revisions October 27, 2023

Approved for publication February 19, 2024