

Building Authentic Science Experiences: Students' Perceptions of Sequential Course-Based Undergraduate Research

Bailey M. Von der Mehden,^{1*} Eric M. Pennino,² Heather L. Fajardo,² Catherine Ishikawa,² and Kelly K. McDonald²

¹Department of Ecology and Evolutionary Biology, The University of Tennessee, Knoxville, TN 37996;

²Department of Biological Sciences, California State University, Sacramento, CA 95819

ABSTRACT

Course-based Undergraduate Research Experiences (CUREs) are attractive solutions for scaling undergraduate research experiences at primarily undergraduate teaching institutions, where resources for faculty research activities can be limited. The Sustainable Interdisciplinary Research to Inspire Undergraduate Success (SIRIUS) project is a unique program that integrates CUREs, coordinated around a local real-world problem, throughout a biology department's curricula. The CUREs are scaffolded to provide all biology majors with multiple opportunities to engage in scientific investigations as they advance through introductory, intermediate, and advanced courses. In this mixed methods, cross-sectional study, we explore students' perceptions of the authenticity of their experiences as they progress through the SIRIUS CUREs. Triangulated data collected from two instruments indicated that students in advanced courses recognized more involvement in research activities and perceived greater authenticity in the science they were performing compared with introductory and intermediate students. Intermediate and advanced students perceived more opportunities for independence; however, experiences with failure and the influence these experiences had on the perceptions of authenticity was primarily observed with advanced students. This study contributes to the growing literature on CUREs with a focus on students from a primarily undergraduate institution with multiple minority-serving designations.

INTRODUCTION

Only half of all students that enter college declaring a Science, Technology, Engineering, or Mathematics (STEM) major graduate with that degree (Chen & Ho, 2012), and this drops to one-fourth for students identifying as Persons Excluded because of their Ethnicity or Race (PEERs; Asai, 2020). One strategy for retaining undergraduate students in STEM disciplines is to provide them with opportunities to participate in experiential learning (e.g., internships, service learning, research). Undergraduate Research Experiences (UREs) have been shown to benefit students in a variety of ways, including enhancing their scientific thinking, science identity, self-efficacy, and in some cases, retention and graduation in their chosen field (Nagda *et al.*, 1998; Seymour *et al.*, 2004; Hunter *et al.*, 2007; Lopatto, 2007; Kuh, 2008; Jones *et al.*, 2010; Eagan *et al.*, 2013; Linn *et al.*, 2015; Robnett *et al.*, 2015).

While UREs offer valuable experiences, access to these opportunities can be limited and exclusive. Student participation in UREs is often dependent on faculty time and resources, students' awareness of research opportunities and benefits, and students' knowledge and confidence in pursuing research positions (Carlone & Johnson, 2007; Bangera & Brownell, 2014; Munawar, 2015; Cooper *et al.*, 2021). In addition to those barriers, UREs are often unpaid and require extra time outside of class, posing additional inequities, especially for those with socioeconomic constraints and nonacademic demands (Bangera & Brownell, 2014). Course-based Undergraduate Research

Erika Offerdahl, *Monitoring Editor*

Submitted Mar 14, 2023; Revised Aug 21, 2023; Accepted Sep 12, 2023

CBE Life Sci Educ December 1, 2023 22:ar46

DOI:10.1187/cbe.23-03-0042

*Address correspondence to: Bailey M. Von der Mehden (bvonderm@vols.utk.edu).

© 2023 B. M. Von der Mehden *et al.*, CBE—Life Sciences Education © 2023 The American Society for Cell Biology. This article is distributed by The American Society for Cell Biology under license from the author(s). It is available to the public under an Attribution–Noncommercial–Share Alike 4.0 Unported Creative Commons License (<http://creativecommons.org/licenses/by-nc-sa/4.0>).

"ASCB®" and "The American Society for Cell Biology®" are registered trademarks of The American Society for Cell Biology.

Experiences (CUREs) represent opportunities for students to participate in research as part of their course curricula (Buchanan & Fisher, 2022). Replacing a prescriptive “cook-book” lab with an authentic investigation (e.g., exploring bacterial diversity in samples collected at different times during a wetlands restoration project) can eliminate many of the barriers students face when accessing faculty-mentored UREs.

Our study used the CURE framework as a theoretical lens to construct our research questions, study design, and analysis strategy. Initially described by Lopatto and colleagues (2008) as “Classroom Undergraduate Research Experiences,” CUREs were further defined by the Course-based Undergraduate Research Experience Network (CUREnet). The latter was established in 2012 to support classroom instruction that integrated five curricular elements: 1) Scientific Practices, 2) Discovery, 3) Relevance, 4) Collaboration and 5) Iteration (Auchincloss *et al.*, 2014; CUREnet). As defined by Auchincloss and colleagues (2014), scientific practices include learning and practicing basic skills/techniques used by scientists (e.g., pipetting, polymerase chain reaction, lab notebook maintenance), as well as applying the process of science (e.g., hypothesis generation, experimental design, data collection/analysis). Discovery refers to the generation of novel data and requires outcomes of student work to be unknown to the students and instructors. Relevance requires that student work is important beyond the classroom (e.g., to the scientific community, local community, or other stakeholders). Collaboration refers to engagement in meaningful research activities with others, such as group experimentation, sharing and discussing ideas, and peer review. Iteration can involve expanding on the work of others or performing an experiment multiple times to troubleshoot issues that arise. The definition of a CURE, and ideas concerning the degree to which each component is emphasized, continues to evolve as the pedagogical strategy is more widely implemented and tested; yet the features described above remain essential to the curricular design.

The benefits of CUREs have been shown to mirror those of UREs (Lopatto *et al.*, 2008). A growing literature cites enhanced self-efficacy, career clarification, and improved retention and graduation rates in students participating in a variety of CURE models (Brownell *et al.*, 2012, 2015; Corwin, Graham *et al.*, 2015; CUREnet; Olimpo *et al.*, 2016; Rodenbusch *et al.*, 2016). For instance, the Freshman Research Initiative, which includes a series of three CUREs for first-year biology students at the University of Texas, Austin, observed an increase in the six-year graduation rate for students who participated (Rodenbusch *et al.*, 2016). Other studies have reported increased content knowledge and scientific thinking (Brownell *et al.*, 2012, 2015), as well as an increased sense of project ownership (Corwin *et al.*, 2018; Cooper *et al.*, 2019), which in turn, has been linked to other psychosocial and attitudinal states that may influence retention (Hanauer *et al.*, 2016). Whereas early investigations predominantly compared CUREs to other laboratory experiences, recent studies have sought to link discrete components or aspects of CUREs to specific student outcomes (Gin *et al.*, 2018; Cooper *et al.*, 2019; Wiggins *et al.*, 2021).

Few studies have explored students’ perceptions of science or research authenticity in their CUREs, and our study aims to expand on this literature. While we primarily draw on the CURE framework established by Auchincloss *et al.*, (2014) for

our definition of research authenticity, we acknowledge the rich literature and variety of definitions that have been applied to the term “authentic” as it is used in science education. Rowland and colleagues (2016) present 26 definitions of “authentic science/research” from a literature review they conducted to construct a framework for the evaluation of their authentic large-scale URE (ALURE) curricula. The authors designed their laboratory courses to emphasize the “real-world” tasks included in the definition of authenticity posed by Jonassen and Rohrer-Murphy (1999), the novel discovery emphasized by Weaver and colleagues (2008), and the concept of flexibility to incorporate product or process-based activities promoted by Spell *et al.*, (2014). However, they also emphasize the importance of collecting students’ perceptions of authenticity to assess the “delivered” curriculum. Furthermore, it has been suggested that authenticity is a dynamic quality of science that relies upon negotiation among those doing science; thus, students working on a project must agree that their project is meaningful, and they must understand for themselves that their work is authentic (Rahm *et al.*, 2003). In a study of high-school students engaging in computer-based and wet lab CUREs, students’ deeply held stereotypes of what scientists do may have influenced them to perceive the wet lab module as more authentic than the database project, despite students recognizing more research tasks performed in the database activity (Munn *et al.*, 2017).

The Laboratory Course Assessment Survey (LCAS) indirectly arrives at student ideas of CURE authenticity by examining the perceptions that students have about participation in activities related to collaboration, novel discovery with broad relevance, and iteration (four of the five CURE curricular components; Corwin, Runyon *et al.*, 2015). Goodwin and colleagues (2021) supplemented the LCAS with an open-ended question that more directly assessed students’ views of authenticity in their comparison of an inquiry-based and CURE laboratory. They found that both student populations recognized they were participating in research-based activities; however, CURE students reported more engagement. Furthermore, by asking students if they thought their laboratory experiences were “real research,” they were able to discern that the CURE curricular elements of scientific practices, discovery with relevance, and iteration, as well as opportunities to fail, were important factors in students’ perceptions of authenticity.

Context of the current study

The Sustainable Interdisciplinary Research to Inspire Undergraduate Success (SIRIUS) project represents one of the first CURE models to be coordinated and implemented extensively throughout a department. The SIRIUS Project adopts a structure in which a series of CUREs are coordinated across a single department’s curricula to address a problem of interest and concern to the local community (i.e., a place-based model; McDonald & Landerholm, 2018; McDonald *et al.*, 2019). SIRIUS was first conceived of, and implemented by, faculty in the Department of Biological Sciences at a minority-serving, comprehensive teaching institution in the western United States. During a multi-year Faculty Learning Community (FLC), CUREs and CURE modules (or mini-CUREs) were designed for 12 courses spanning introductory through capstone laboratories with research activities focused on the

TABLE 1. SIRIUS Project CURE levels and courses

Course level	Course title	Number of students per semester	Research topic
Introductory	BIO 1: Biodiversity, ecology and evolution	~384	Fertilizer effects on American River macroinvertebrates
	BIO 2: Cells, molecules and genes	~142	Urban run-off effects on American River soil bacteria
Intermediate	BIO 139: General microbiology	~142	Identification of fecal coliform bacteria from the American River
	BIO 160: General ecology	~72	Ecology of American River tree root fungi
Advanced	BIO 127: Developmental biology	~48 students per year (Fall semesters only)	Human-derived toxin effects on population dynamics of <i>C. elegans</i> nematodes
	BIO 128: Plant physiology	~24 students per year (Spring semesters only)	Effects of restoration adjacent to the American River on mycorrhizal fungi diversity
	BIO 145: Diversity of microorganisms	~16 students per year (Spring semesters only)	Identification and sequencing of American River antibiotic-producing bacteria (Tiny Earth)
	BIO 180: Advanced molecular biology	~32 students per year (Spring semesters only)	Genetic engineering of American River soil bacterium <i>Acinetobacter baylyi</i>
	BIO 187: Advanced cell biology	~16 students per year (Spring semesters only)	Analysis of the effects of toxins on the survival of mammalian cell cultures

health of, and human impacts on, the American River and its surrounding ecosystem.

The overarching goals of the SIRIUS Project are to 1) provide research opportunities to all biology students at the institution, 2) develop a scientific community where students, faculty, and staff work collaboratively on a local problem, and 3) scaffold curricula to allow students multiple opportunities to build their research skills and identity (McDonald & Landerholm, 2018; McDonald *et al.*, 2019). At the time of graduation, biology majors have generally participated in four to eight SIRIUS Project courses, thus performing research across multiple sub-disciplines of biology. As of 2020, over 5000 students participated in SIRIUS courses and a comprehensive evaluation plan to study the impacts of the program has assessed a variety of student and faculty outcomes (McDonald *et al.*, 2019; Martin *et al.*, 2021).

This widespread implementation allowed us a unique opportunity to observe how student perceptions of whether they were doing real (“authentic”) science differed after completing CUREs in courses at varying levels (introductory to advanced), as well as compare the rationales for their perceptions. While the study was cross-sectional, and therefore changes in individual students or cohorts could not be tracked, students in the advanced classes would have had several CURE experiences before taking the advanced class. Thus, at least some observed outcomes could be due to the cumulative effect of experiencing multiple related CUREs. Here, we present data from an examination of the perceptions students hold about the authenticity of their lab experiences at three timepoints (introductory, intermediate, and advanced). Specifically, we address the following research questions:

1. Do students recognize participating in research-based activities in their laboratory curricula?
2. Do students perceive their laboratory experiences to represent authentic science? Why or why not?
3. Do perceptions and rationales of authentic science differ between students enrolled in introductory, intermediate, and advanced CUREs or across different demographic groups?

METHODS

SIRIUS courses

The courses included in this study spanned the curricula and contributed to the requirements or electives of all concentrations offered by the Department of Biological Science (i.e., General BA and BS; Biomedical Sciences; Cell and Molecular Biology; Clinical Laboratory Sciences; Ecology, Evolution and Conservation; Forensic Biology and Microbiology). The research projects differed between individual courses (Table 1), but all curricula related to the health of and/or human impacts on the American River, which runs through campus and is listed as Impaired under Section 303(d) of the federal Clean Water Act of 1972. As described in the following section, SIRIUS courses were categorized according to introductory, intermediate, and advanced levels. All courses included in the SIRIUS Project have lecture and lab components, and CURE curricula were integrated into laboratory sections (and in the case of BIO1, a discussion section). Only the nine SIRIUS laboratory courses that met the definition of a CURE were evaluated in this study. We used syllabi, course materials, and discussions with faculty to ensure that their CURE engaged students in scientific practices, involved collaboration and iteration, and addressed a novel question with relevance outside their specific course (Fajardo, 2022).

CURE levels

SIRIUS Project CURE courses were assigned to introductory, intermediate, and advanced levels (Table 1) based on a combination of three parameters: 1) the year in a student's undergraduate career the courses are typically taken (as first-year student, Sophomore, Junior, Senior or Graduating Senior), 2) the number of Biology and Chemistry prerequisites required by the course, and 3) the nature of the curricula; in particular, the length of research project and the inclusion of central versus peripheral tasks (Table 2). Central tasks are fundamental to the scientific method and may involve developing a research question and/or experimental design (Corwin, Graham *et al.*, 2015; Dolan & Weaver, 2021). Peripheral tasks are those that a novice

TABLE 2. Characteristics of SIRIUS Project CUREs and instructors by level and course

Level	Course	Year taken	Cumulative Prerequisites (BIO and CHEM)	Length of CURE	Instructor type	FLC Participant
Introductory	Bio 1	First Year/So, non-majors	none	30% of semester	TT, LTL, L, GTA	TT, GTAs
	Bio 2	So/Jr	2	75% semester, with skills building included	TT, LTL, L, GTAs	TT, L
Intermediate	Bio 139	Jr/Sr	6	30% actual experiment but prior weeks for skills development necessary	TT, LTL, L	TT, LTL
	Bio160	Jr/Sr	5	Full semester	TT, LTL, L	TT, LTL
Advanced	Bio127	Sr/Graduating Sr	7	50% for research question, but prior weeks for skills development	TT	TT
	Bio128	Sr/Graduating Sr	4	Full semester	TT	TT
	Bio145	Graduating Sr	7	Full semester	TT	TT
	Bio180	Graduating Sr	6	Full semester	TT	TT
	Bio187	Graduating Sr	6	Full semester	TT	TT

Student years are defined as follows: First Year (1–29.9 semester units); So (Sophomores, 30–59.9 units); Jr (Juniors, 60–89.9 units); Sr (Senior, > 90 units); Graduating Sr (Students scheduled to graduate within the year).

Sr and Graduating Sr are separated due to the low four-year graduation rate and the tendency of students to accumulate more than 120 units by the time of graduation. Instructors are categorized as tenure track (TT); long-term lecturer (LTL, with three-year contracts); lecturer (L, with semester or year-long contracts); and graduate teaching associate (GTA, with one semester of pedagogical and practical training for the courses they teach).

could engage in to learn and contribute to an experiment and are also tasks that allow for mistakes and repetition, such as data collection and analysis (Lave & Wenger, 1991; Corwin, Graham *et al.*, 2015; Dolan & Weaver, 2021). The introductory courses are the first biology courses taken by majors (and some nonmajors) and have few or no prerequisites. Students in introductory courses learned and practiced peripheral tasks (e.g., calculations and pipetting) at the beginning of the semester, then proceeded to more advanced peripheral tasks (e.g., collecting and analyzing data) in the next stage of their project; however, the research questions were defined for them. In one of the introductory courses, students performed central tasks when revising experimental designs for a second iteration of experiments. Introductory courses are prerequisites for intermediate courses, which are typically taken in years three and four. Intermediate courses built on skills learned at the introductory level; however, activities included more central tasks, such as making choices regarding strategies and tests that should be performed given a specific experimental objective. Advanced courses have the most prerequisites and are often composed of students who have senior class status and are nearing graduation. Advanced courses incorporated higher levels of inquiry (Buck *et al.*, 2008), with CURE projects that were typically longer and allowed students more autonomy in decision making. Advanced CURE students engaged in more central tasks compared with the other levels, such as generating research questions, developing proposals, and experiments.

In addition to curricular differences, faculty type and involvement in the SIRIUS professional development activities differed between levels (Table 2). In advanced classes, smaller enrollments allowed single instructors to teach all aspects of the course. These instructors all participated in the SIRIUS professional development. In the larger introductory and intermediate classes, faculty participating in SIRIUS designed the CUREs, but had less control over the implementation of the curricula in

the multiple laboratory sections taught by lecturers. All instructors teaching CUREs at the introductory and intermediate levels were aware of the project, but not all had the same degree of training or level of investment.

Assessment instruments

Laboratory course assessment survey (LCAS). The 17-item Laboratory Course Assessment Survey (LCAS) was used to measure students' perceptions of laboratory course curricula across the three CURE levels for three independent constructs: Collaboration, Discovery/Relevance, and Iteration (Corwin, Graham *et al.*, 2015, Supplemental 1). The Collaboration section of the LCAS asked students how often they engaged in six collaborative activities with possible total scores ranging from 6–30. Example activities in the Collaboration section are “discussing elements of their investigation with classmates or instructors,” “helping other students collect or analyze data,” and “providing constructive criticism to classmates.” The Discovery/Relevance construct of the LCAS consisted of five items that prompted students to rank their level of agreement that they participated in activities such as “conducting an investigation to find something previously unknown,” and “developing new arguments based on data.” Possible total scores in this section ranged from 5 to 30. The Iteration section asked students to indicate their level of perceived agreement that they participated in six activities including, “revising or repeating work to account for errors or fix problems,” “having time to share and compare data with other students,” and “having time to revise drafts of papers or presentations about my investigation based on feedback.” Possible total scores in the Iteration section ranged from 6 to 36.

Authenticity Questions. Along with the LCAS, we posed two free-response questions to gather students' ideas about the nature of science and their participation in the process of science (Supplementary 1 and 2). Using a phenomenological

TABLE 3. Data collection schedule

Assessment instrument	Level	Course	Semester of data collection
1. Laboratory Course Assessment Survey (LCAS)	Introductory	Bio 1	F15, S16, S17
		Bio 2	F15, S16, S17
	Intermediate	Bio 139	F16, S17
		Bio 160	F16, S17
2. Authenticity Question	Advanced	Bio 127	F16
		Bio 128	S17
		Bio 145	S18
		Bio 180	F18
		Bio 187	S18, S19

"F" and "S" indicate Fall and Spring semesters.

approach to understand the CURE experiences from the participants' perspective (Easterby-Smith *et al.*, 2012), we asked students the questions: "Did you feel like you were doing 'real science' in your lab course? Why or Why not?"

Study population and data collection

This study surveyed students enrolled in SIRIUS Project courses at a primarily undergraduate serving institution in the western United States (Table 1) under approved institutional review board protocol #13-14-148. All students enrolled in SIRIUS courses during semesters under evaluation (Fall 2015–Spring 2019) were recruited by their instructors during class to participate in surveys (Table 3). A small number of students appeared in the overall dataset more than once, as they participated in multiple SIRIUS courses during the evaluation timeframe. Survey items were specific to the courses students participated in, and thus were included in the data set.

The assessment instruments were administered as post-CURE surveys at the end of each semester under evaluation (Table 3). Surveys conducted in early semesters (2015–2016) were administered in class on paper. Surveys conducted in subsequent semesters (2017–2019) were administered outside of class using the Campus Labs online survey tool.

Data analysis

Demographic data. Demographic data were obtained from the university's Office of Institutional Research and variables were summarized using descriptive statistics. Because our analyses involve comparisons across levels, χ^2 tests of independence were used to examine the composition of students in each CURE level according to the following demographic groups: gender (Female, Male), race/ethnicity (Asian; Persons Excluded due to Ethnicity or Race, which included Latinx, Black, Native American, and Pacific Islander; White), Pell Grant eligibility status (Pell Grant eligible, Non-Pell Grant eligible) and transfer status (Transfer, Nontransfer). At the time of data collection, the office of institutional research only offered two options for "gender identification" (male and female). While students self-reported this information on admission documents, the limited selection options mean we were unlikely to capture the true gender diversity of our study participants. Also, the term "transfer" in this study refers to any college student transferring from another institution (community college or four-year) to our institution; however, the vast majority of transfer students at our institution (85% or more) are from community colleges.

Quantitative analysis of Laboratory Course Assessment Survey (LCAS). Student responses for items in each construct (Collaboration, Discovery/Relevance, and Iteration) were examined and eliminated if they met one of three criteria: at least one blank item, at least one response of "I don't know," or at least one response of "I prefer not to respond." The scale for the Collaboration construct was flipped before analysis to be more consistent with the other two constructs. Counts from items in each construct were summed to provide a single score for each student. The LCAS has been validated by prior studies across large undergraduate student populations (Corwin, Runyon *et al.*, 2015; Corwin *et al.*, 2018; Cooper *et al.*, 2019; Goodwin *et al.*, 2021). On our student population, Confirmatory Factor Analysis indicated reasonable fit (CFI = 93.6, TLI = 92.5, RMSEA = 0.066) reasonable fit is defined by CFI and TLI > 90.0 and RMSEA between 0.08 and 0.05 (Bentler, 1990; Byrne, 1994; Fan *et al.*, 1999). Cronbach's alpha internal consistency reliabilities for each construct Collaboration ($\alpha = 0.86$), Discovery/Relevance ($\alpha = 0.85$), and Iteration ($\alpha = 0.89$), were above the accepted cut off ($\alpha = 0.70$) (Nunnally, 1978). All statistical tests were performed in R v. 4.0.3 (R Core Team, 2021) using the following packages: psych (Revelle, 2022), ltm (Rizopoulos, 2006), lavaan (Rosseel, 2012), dunn.test (Dinno, 2017), and ggplot2 (Wickham, 2016).

To determine if we could combine courses into introductory, intermediate, and advanced levels for LCAS analyses, we compared courses within each level using ANOVA. We found no significant differences between mean scores in courses within the same level, thus the courses were combined into the three categories for further analysis. Preliminary analyses also indicated nonnormality and nonhomogenous variances. Therefore, nonparametric Kruskal-Wallis tests were used to compare median LCAS scores across levels (introductory, intermediate, and advanced) for each construct. Post hoc Dunn tests with the Benjamini-Hochberg adjustment were used to determine which levels were significantly different from each other (Dinno, 2015).

Qualitative analysis of Authenticity Question responses. Student excerpts ($n = 1,252$) from the Authenticity Question were managed and analyzed using Dedoose version 9.0.17 (Dedoose, 2021) and Microsoft Excel (Microsoft Corporation, 2018). A preliminary version of the codebook was developed for a subset of the data by one of the authors (K.M.) and a former researcher. In the preliminary version, analysis of the responses

TABLE 4. Student demographics of SIRIUS course levels

Demographics	Introductory <i>n</i> = 746%	Intermediate <i>n</i> = 406%	Advanced <i>n</i> = 100%
Gender			
Female	60.7	57.9	60.0
Male	37.8	39.7	39.0
Unknown	1.3	2.5	1.0
Race/Ethnicity			
Asian	29.4	32.0	26.0
PEERs	33.8	27.1	28.0
White	26.7	28.8	27.0
Other	6.8	8.1	14.0
Unknown	3.4	3.9	5.0
Pell status			
Pell eligible	54.4	50.0	56.0
Non-Pell eligible	43.2	46.3	42.0
Unknown	2.4	3.7	2.0
Transfer status*			
Transfer	23.9	54.7	52.0
Nontransfer	74.7	42.6	46.0
Unknown	1.5	2.7	2.0

Chi-square tests were used to examine composition of sociodemographic variable across levels.

*Indicates significance ($p < 0.001$).

identified: 1) a priori codes that directly represented the CURE curricular elements drawn from the CURE framework (i.e., Scientific Practices, Discovery, Relevance, Collaboration, and Iteration), and 2) emergent codes, which were reoccurring words, phrases, and ideas in student responses. The first codebook, amended by three researchers (B.V., E.P., and H.F.), was used for analysis of a larger data set for this study (Supplementary 3). To develop the new codebook, descriptive coding methods were used to define emergent codes that summarized, in one or two words, the main ideas from students' responses (Saldana, 2021). The codebook was further adapted by adding, revising, and removing codes as necessary during the iterative coding process, which used an additional subset of the data. For this study, two or more researchers independently applied codes to all student excerpts. Initial percent agreement was 75.7% or greater, and all coding disagreements were resolved by discussion to reach consensus. Representative excerpts were selected to illustrate code usage across introductory, intermediate, and advanced levels. A priori and emergent code frequencies were evaluated for differences across demographic groups (gender, race/ethnicity, Pell Grant eligibility status, and transfer status). Descriptive statistics were used to

summarize frequency of Yes/No responses and rationales for Yes/No responses (rationale codes).

RESULTS

Demographic characteristics of students across course levels

From Fall 2015 to Spring 2019, 1115 unique students out of 1252 total student participants completed the surveys, with a majority identifying as female (59.8%), students of color (61.3%), and eligible for Pell grants (53.1%). Demographic data were analyzed to determine the student composition of each SIRIUS CURE level (Table 4). Transfer status was the only factor that differed across the course levels ($p < 0.001$), with significantly fewer transfer students in the introductory courses than the intermediate and advanced levels. This is due to most transfer students completing their introductory coursework at community colleges.

Student perceptions of collaboration, discovery/relevance, and iteration

The Laboratory Course Assessment Survey (LCAS) was used to measure students' perceived participation in activities related to Collaboration, Discovery/Relevance, and Iteration in their SIRIUS CURE courses. For all constructs, higher scores indicated a higher level of perceived participation. Kruskal-Wallis tests indicated significant differences between course levels for each construct (Collaboration, Discovery/Relevance, and Iteration; Table 5).

Response rates for the LCAS portion of the survey were 63.6% for introductory, 72.5% for intermediate, and 66.5% for advanced students. Students in all CURE levels had relatively high Collaboration scores (Figure 1); thus, they reported high levels of participation in activities such as group work and providing feedback to peers. Post hoc Dunn tests indicated advanced students reported participating in collaborative activities more often than introductory and intermediate students ($p < 0.001$). Although the medians are the same between introductory and intermediate groups, the two groups were statistically different ($p < 0.05$).

The Discovery/Relevance construct of the LCAS prompted students to rank their level of agreement that they participated in activities related to novel discovery with broad relevance. Advanced students had the highest Discovery/Relevance scores rating participation in these activities more often than introductory and intermediate students ($p < 0.001$; Figure 1).

Differences were reported across all course levels for laboratory activities related to Iteration, such as troubleshooting, repeating and revising work. Advanced students had higher Iteration scores than introductory students, who had higher scores

TABLE 5. Possible scores for each construct are represented by score ranges

Construct	Score Range	Introductory		Intermediate		Advanced		Kruskal–Wallis		
		Median	IQR	Median	IQR	Median	IQR	<i>p</i>	<i>h</i>	<i>df</i>
Collaboration	6–30	27	23–30	27	22–30	30	28–30	$0 < 0.001$	37.5	2
Discovery/Relevance	5–30	24	21–26	23	19–26	28	26–30	$0 < 0.001$	103	2
Iteration	6–36	28	24–31	26	20–30	33	30–36	$0 < 0.001$	114	2

Each construct has a different number of items, therefore, score ranges vary. Kruskal–Wallis results are to the right of each construct.

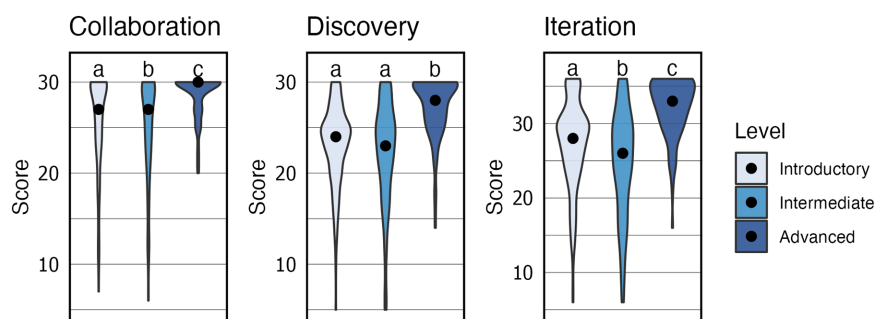


FIGURE 1. Comparisons of LCAS Scores Between CURE Levels. Each construct had a different number of participants due to the exclusion criteria used to prepare data for analysis. Collaboration LCAS scores from CURE course levels: introductory ($n = 579$), intermediate ($n = 342$), and advanced ($n = 95$). Discovery/Relevance scores from CURE course levels: introductory ($n = 627$), intermediate ($n = 337$), and advanced ($n = 105$). Iteration scores from CURE course levels: introductory ($n = 652$), intermediate ($n = 326$), and advanced ($n = 103$). Black dots represent medians. Levels with the same letter are not significantly different ($p > 0.05$).

than intermediate students, and post hoc Dunn tests indicated all scores were significantly different from one another ($p < 0.001$; Figure 1). Overall, the LCAS data indicated that students in every level recognized participating in research-based activities, with advanced students perceiving the most participation with the least score variation (Figure 1 and Table 5).

Student perceptions of science authenticity

A total of 1252 students responded to the survey question “Did you feel like you were doing ‘real science’ in your lab course? Why or why not?” Responses were collected from 746 introductory, 406 intermediate, and 100 advanced students. Response rates for the open-ended authenticity questions were 61.2% for introductory, 79.8% for intermediate, and 63.3% for advanced students. Data collected from the first question, coded as Yes, No or Maybe, indicated the majority of students perceived their laboratory experiences to be “real science.” Eighty-seven percent of students across all levels combined responded positively (i.e., Yes), and as course level increased, the proportion agreeing to the statement increased, with nearly all (98%) of advanced students indicating their laboratory experiences represented “real science.” Advanced students were also more decisive in their responses, as evidenced by the lack of Maybe/Unclear responses (Table 6). The second question, “Why or why not?” offered students the opportunity to provide the reason for their perceptions of authenticity or inauthenticity. Each student excerpt was first coded with the a priori CURE curricular elements, as in the following example from an advanced student whose response included activities or ideas

TABLE 6. Student responses to the Authenticity Question

Responses	Introductory $n = 746\%$ (n)	Intermediate $n = 406\%$ (n)	Advanced $n = 100\%$ (n)
Yes	85.0 (634)	89.4 (363)	98.0 (98)
No	6.8 (51)	6.2 (25)	2.0 (2)
Maybe and Unclear	8.3 (61)	4.4 (18)	0.0 (0)

“Did you feel like you were doing ‘real science’ in your lab course?”

related to Relevance, Original/Discovery, and Iteration.

“Yes. Because we were working with genes in a system that had not been tested before and investigating a topic that needs more research on. Our research can be used by other scientists in future work to continue understanding gene amplification. Also, the outcome was not known so we were doing novel research.”

Quantifying the CURE curricular elements revealed that the responses from advanced students contained more CURE curricular elements (per excerpt) than responses from introductory and intermediate students (Table 7). The majority of students across all levels mentioned one CURE curricular element in their responses; however, the percentage of students within each level mentioning two to four CURE curricular elements increased

with course level. Across all CURE levels, participation in Scientific Practices (>59%, Figure 2) was cited most often as the reason students perceived their lab experiences to be authentic, while Collaboration was mentioned the least ($\leq 5\%$). Advanced students cited Relevance, Original/Discovery, and Iteration more often than introductory and intermediate students, who reported these reasons at similar frequencies (Figure 2). After Scientific Practices, advanced students mentioned Original/Discovery most often, followed by Iteration, and then Relevance. Intermediate and introductory students both mentioned Iteration as their rationale least often with Relevance cited slightly more often than Original/Discovery.

Scientific Practices was cited most often by students at all course levels, and due to the broad nature of the code, we performed additional open, descriptive coding to disaggregate it into the following (emergent) sub-codes: Techniques, Answering Questions, Communication, Analysis, Hypothesis Creation, and Fieldwork. Performing various techniques was the most frequently discussed subcode (87.2%) of Scientific Practices across all levels (Figure 3). Advanced students mentioned Answering Questions and Communication more often than students at the other levels; however, all codes other than Techniques were reported infrequently by comparison.

Additional emergent codes, identified through inductive open coding, provided further insight into student perceptions

TABLE 7. CURE curricular element (collaboration, iteration, discovery/relevance) mentions per student response

# Of CURE curricular elements per excerpt	Introductory $n = 746\%$ (n)	Intermediate $n = 406\%$ (n)	Advanced $n = 100\%$ (n)
0	22.1 (165)	22.9 (93)	10.0 (10)
1	61.7 (460)	59.1 (240)	45.0 (45)
2	14.6 (109)	16.0 (65)	33.0 (33)
3	1.5 (11)	2.0 (8)	11.0 (11)
4	0.1 (1)	0.0 (0)	1.0 (1)

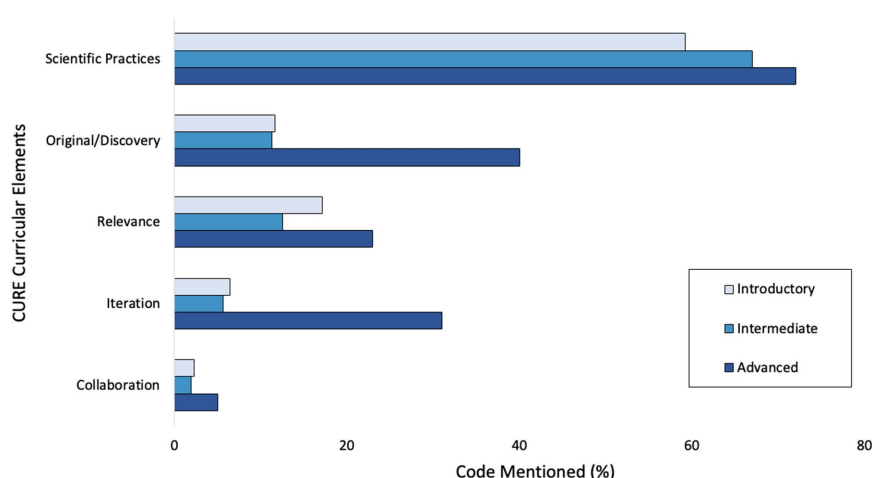


FIGURE 2. CURE curricular elements from authenticity question compared across levels. CURE curricular elements (a priori rationales) across introductory ($n = 746$), intermediate ($n = 406$), and advanced ($n = 100$) students.

of their laboratory experiences. While all emergent codes were cited less frequently than the CURE curricular elements, advanced students referred to Independence (17.0%), and Resilience/Failure (14.0%) more often than introductory and intermediate students (Figure 4), and intermediate students mentioned experiencing Independence (13.5%) and Resilience and Failure (3.20%) more often than introductory students (4.29%, 0.27% respectively), which may point to increased rigor across the scaffolded laboratory curricula.

Five codes emerged from responses of students who did not think their labs represented “real science”: Limited, Cookbook, No Structure, No Engagement and No Relevance. Although these negative emergent codes were cited infrequently (<9%) by students in all levels, introductory students were more likely to question their laboratory curricula’s inauthenticity for these reasons (Figure 5). The most common code for this group was

“Limited,” which we applied when students mentioned feeling their experience was lacking something, such as time to perform a certain task, involvement in study design, or access to materials necessary for experimentation. Limited and Cookbook were the only two negative codes to emerge from advanced student responses, whereas a few introductory and intermediate students also indicated that their labs lacked structure or relevance or were not engaging. Example quotes for the a priori CURE curricular element codes and the most frequently reported emergent codes are provided in Table 8. These quotes are representative of the code descriptions and serve to illustrate the differences in student responses across CURE levels.

DISCUSSION

The higher discovery/relevance and iteration LCAS scores from our advanced classes, relative to lower division courses, are consistent with what we would expect from a curriculum of scaffolded CUREs. Scores from our advanced classes were higher in the discovery/relevance and iteration LCAS constructs compared with scores from our lower division courses, and also higher than other courses reported in the literature (Table 9; Beck *et al.*, 2023). This is especially notable given that the results were not from a single class, but rather from five different classes with five different instructors. Several factors may have contributed to these higher scores. Smaller class sizes and students’ prior knowledge and skills could have allowed faculty to genuinely embed a greater amount of authentic work in the course. Sathy *et al.*, (2020) also observed higher scores in their advanced research methods for a psychology course (taught by the same instructor as their introductory research methods course) and suggested smaller class sizes as a possible cause. An additional

factor might be consistent messaging from faculty teaching at the advanced level, all of whom participated in the SIRIUS FLC professional development activities. By contrast, lower medians and greater variability in scores for intermediate classes may have been due to the fact that fewer of the lab instructors for these courses participated in the SIRIUS FLC. A third possible factor is that students in the advanced classes may have previously participated in introductory or intermediate level SIRIUS CUREs, and repeated experiences with CUREs might have helped students recognize when they were engaging in relevant discovery and iteration.

The lower LCAS scores for iteration observed in intermediate classes compared with introductory and advanced courses underscore the value of LCAS scores as a tool for assessing individual course and program impacts. The lower

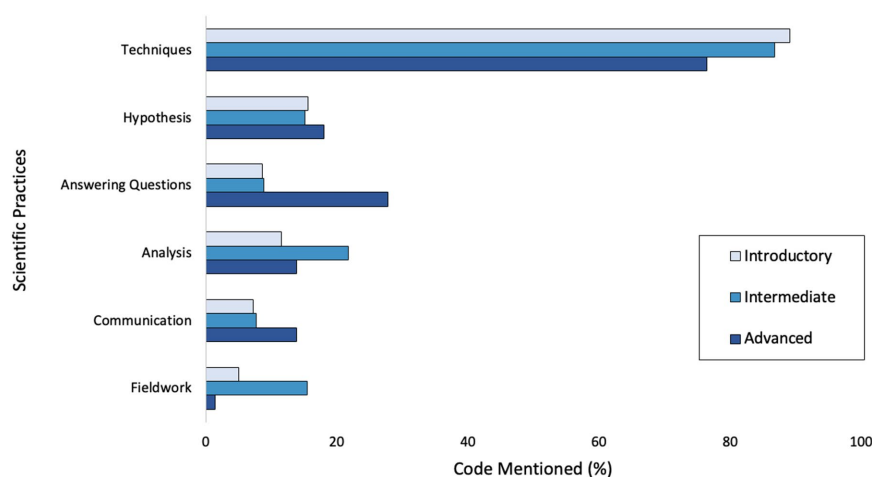


FIGURE 3. Scientific practices disaggregated and compared across levels. Subcodes nested under scientific practices compared across introductory ($n = 442$), intermediate ($n = 272$), and advanced ($n = 72$) students.

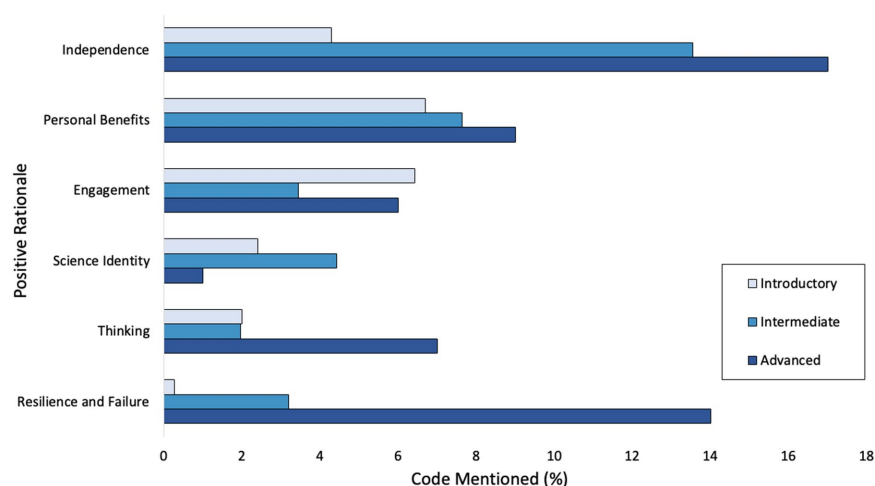


FIGURE 4. Emergent positive rationales compared across levels. Codes are compared across introductory ($n = 746$), intermediate ($n = 406$), and advanced ($n = 100$) students.

iteration scores in intermediate classes were consistent with the CURE designs course instructors described to the project coordinators. Introductory and advanced CUREs were typically longer in duration and provided more opportunities for repetition and revision than intermediate CUREs. This agreement supports Beck and co-authors' assertion that the utility of the LCAS lies in helping us identify what students are doing in their laboratory classes (Beck *et al.*, 2023).

Interestingly, however, iteration did not appear to contribute strongly to introductory- or intermediate-level students' perceptions that they were doing real science, based on answers to the open-ended question. While an actual lack of iteration could explain these results for intermediate-level courses, introductory-level courses experienced iteration but did not mention it as an aspect of real science. Advanced-level courses also experienced iteration in their CUREs and students mentioned

are the CURE elements (as defined in Auchincloss 2014) most likely to contribute to students' perception that they are doing real science.

Student answers to our open-ended question about whether they did real science in their courses also suggest that students in advanced courses were not only experiencing more sophisticated CURE curricula, but also developing broader views about what it means to do authentic science. Consistent with other studies of student and faculty perceptions of what constitutes authentic science (Spell *et al.*, 2014; Rowland *et al.*, 2016; Wiggins *et al.*, 2021), scientific practices were the most frequent reason provided by students at any level. When disaggregating scientific practice in our study, however, students from advanced classes more frequently referenced practices that involved scientific thinking (as opposed to physical lab or field techniques) than students from intermediate and introductory

courses. This observation is consistent with curriculum analysis of SIRIUS courses at the three levels, which indicated that students engaged in more peripheral tasks at the introductory- and intermediate-level and more central tasks in advanced-level courses (Fajardo, 2022). While students at all levels were participating in real research, the degree of participation was intentionally scaffolded, so students progressed into deeper cognitive involvement in the research activities as they progressed through the curriculum.

Disaggregating relevance and discovery in open-ended responses revealed that students' perceptions of relevance were mostly consistent across levels, but students from the advanced classes mentioned discovery nearly four times as often as students in introductory or intermediate classes. Research in each course had an element of novelty, but the degree may

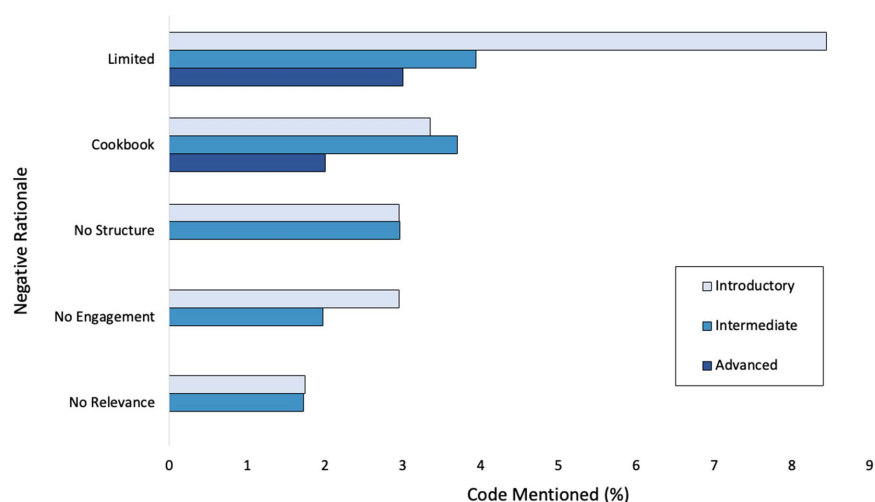


FIGURE 5. Negative rationales compared across levels. Codes are compared across introductory ($n = 746$), intermediate ($n = 406$), and advanced ($n = 100$) students.

TABLE 8. Examples of coded student responses

Authenticity code	Definition	Introductory quote	Intermediate quote	Advanced quote
Scientific practices	Using the scientific method, including skills or tools used to do so	Yes, it was my first time doing an experiment on site rather than just using the samples and solutions given the lab room.	Yes. We took samples from the American River and recorded our data.	Yes. I collected the sample and processed everything through as well as plating, doing physiological test and extracting chemicals from bacteria.
Original/Discovery	Finding novel results, the results were previously unknown	Yes, because we received the opportunity to investigate a hypothesis that nobody knows the answer to yet.	Yes. We were doing an experiment that attempted to answer question that hadn't been answered before.	Yes. We were acquainted with the scientific method, but the river ecosystem had unknown results.
Relevance	References some sort of greater purpose, contributing to knowledge in some way	Yes. I felt that the work I was doing was actually determining real variables that affect the American River.	Yes. We dealt with real issues that affect the world today. Such as antibiotic resistance and common bacteria.	Yes. The American River is a big contributor in Sacramento. Understanding what is going on in the river and the ecosystem will benefit the community. We will be more responsible in how we use it.
Iteration	References growth, building on the work of others, or repetition	Yes. because we were given the opportunity to do trials for our experiment	Yes. There was a trial and error for our lab.	Yes. Yes, not all of the experiments generated an expected result. Troubleshooting was always a major skill to look at what worked or didn't work to proceed with the experiment.
Collaboration	Working with peers, faculty, staff	Yes, because I got to do experiments with my peers, we collected the data and analyzed it.	Yes, a group of my peers and I conducted our own experiment.	Yes. We were expected to... learn how to work in a group and be able to switch around working in the lab.
Independence	Working on experiment by themselves	Yes, because we are finding our own results.	Yes. This was the first class that I really had to conduct my own experiment from scratch.	Yes. I am ...forming my own hypothesis. I am also interpreting my data, and these are not results that the professor has already set up.
Personal Benefits	Gaining content knowledge or skills that will be beneficial in later classes or careers	Yes. Lab constantly challenges me to further my knowledge. It teaches me hands on skills that I would never have learned without lab.	Yes. I think what we did ... helped us to build a stronger foundation for more in-depth biology classes to come.	Yes. ...Also, I was able to learn new lab techniques that are crucial for a career.
Resilience and Failure	References an experiment not working but understanding that is part of science	Yes...Although the results did not come out the way we wanted, we can still learn from it and our mistakes.	Yes. because none of my results were ever correct or neat on the first try. I experienced a lot of problems a scientist would come across in a lab setting.	Yes. This lab course was challenging and informative as it has taught me to be confident in myself even though I was not achieving the goals I wanted, but it taught me to adapt and solve problems and then results aligned with those goals.
Limited	Experience was missing something the student felt should have been included	Not necessarily because our only job was to investigate the species' richness.	Maybe. The time and resource constraints were difficult to dedicate to a hypothesis in depth.	Yes. To some degree. I would have liked to be involved in the process of plating the bacteria and worms.

The first five rows are examples of a priori coded responses across levels.

Some quotes are excerpted sections of entire student responses to provide clear examples of the authenticity code.

TABLE 9. LCAS scores reported in the literature for a variety of disciplines and course levels

Reference	Course type	Percent of range		
		Collaboration	Discovery	Iteration
Corwin, Runyon <i>et al.</i> , (2015)	Biology courses with CUREs (intro and upper division)	88	81	80
Corwin, Runyon <i>et al.</i> , (2015)	Biology courses without CUREs (intro and upper division)	87	69	74
Allen <i>et al.</i> , (2021)	Chemistry (upper division, Analytical Chemistry)	92	81	81
Sathy <i>et al.</i> , (2020)	Introductory Statistics	83	79	76
Sathy <i>et al.</i> , (2020)	Research Methods (Psychology)	76	85	81
Sathy <i>et al.</i> , (2020)	Advanced Research Methods (Psychology)	88	91	85
Cruz <i>et al.</i> , (2020)	Organic Chemistry (large enrollment, lower division)	72	74	72
Kearney (2022)	Introductory Biology	90 ^a	85 ^a	78 ^a
Goodwin <i>et al.</i> , (2021)	Introductory Biology–Inquiry	81	77	74
Goodwin <i>et al.</i> , (2021)	Introductory Biology–CURE	87	87	88
This study	Introductory Biology courses	90	80	77
This study	Intermediate-level Biology courses	87	77	72
This study	Advanced Biology courses	100	93	92

^aEstimated from graph.

Percentage of the possible range is shown because some studies used slightly modified versions of the survey, resulting in different high and low scores. The numbers in each column represent average scores of the constructs, represented as a percentage of the total range of the scores reported.

have differed. For example, BIO 1 students manipulated fertilizer levels in microcosms filled with water and sediment from the river by campus. While the instructor did not know what to expect with the faunal assemblages in these specific samples, students may have suspected that this was something that had been done before in other rivers. Also, instructor communication may have played a role, with advanced-level instructors placing greater emphasis on the novelty of their work than introductory or intermediate instructors. Interestingly, for our advanced students, the idea that they were doing something “novel” or investigating a question that no one had answered appears to have supported their work’s authenticity more than the potential for their work to be meaningful to a broader community. This result suggests that considering relevance and discovery as distinct items (as opposed to combining them, as the LCAS does) may be useful when modeling and testing positive effects of CURE elements.

Our study found that students commented more about failure in the advanced courses than the introductory or intermediate courses. Failure and frustration, with the opportunity to troubleshoot, have emerged as useful characteristics of CUREs (Lopatto *et al.*, 2020), as well as reasons students perceive their CURE curricula to represent authentic research (Rowland *et al.*, 2016; Goodwin *et al.*, 2021; Wiggins *et al.*, 2021). Both Goodwin *et al.*, (2021) and our study saw examples of students saying the research was “real” because it didn’t work. Unexpected challenges lend a sense of authenticity and create a contrast with “cookbook” labs that have been designed to reliably illustrate certain outcomes. Another aspect of failure that appeared in our study and others is students’ realization that failure is a part of real research—something they will encounter in the future (Rowland *et al.*, 2016; Goodwin *et al.*, 2021). Our results support those of a number of studies that have found that students value the opportunity to fail, troubleshoot, and repeat an experiment (Rowland *et al.*, 2016; Lopatto *et al.*, 2020; Wiggins *et al.*, 2021). Students recognize that this cycle helps them learn, is valuable preparation for future research positions, and can be enjoyable when it leads to success. Conversely, a small number of students felt like it was not real science if they did

not have time to repeat their work until they were successful in both our study and Goodwin and colleagues’ (2021). Overall, however, the ability to fail and repeat appears to be a positive benefit of CUREs, especially in advanced courses.

Students’ overall perception that they were doing real science in the SIRIUS Project CUREs, regardless of their rationales, is encouraging and important for several reasons. Research suggests that student perceptions of specific CURE components correlate with desired student outcomes. For instance, studies have shown that perceptions of relevance and discovery (as measured by the LCAS) are positively correlated with a sense of project ownership (Corwin *et al.*, 2018; Cooper *et al.*, 2019). Cooper *et al.*, (2019) used the LCAS to evaluate two sections of a course that differed only in the degree of relevance and novel discovery and found that students perceived this difference and also reported greater project ownership in the section with relevance and discovery in the design. A sense of project ownership has been associated with outcomes such as persistence in science (Hanauer *et al.*, 2016), suggesting the inclusion of these CURE components could have both short and long-term influences. Finally, iteration and failure can help students develop stronger experimental design skills even in laboratory courses that lack strong relevance and discovery components (Lansverk *et al.*, 2020).

In addition to the psychosocial and cognitive benefits described above, students who recognize the authenticity or importance of their work may be more likely to highlight it when pursuing future academic or career opportunities. Some faculty in advanced SIRIUS courses explicitly addressed how students might include their CURE research in resumes, graduate school applications, or interviews, thus providing a more diverse student population with access to advanced academic and career paths. Another potential benefit that could be explored is that understanding characteristics of real research (like failure and iteration) might ease the transition into graduate work for students, through changing expectations and building resilience.

Limitations

The greatest limitations of this study involved data collection and sample sizes, which also influenced the application of a

cross-sectional, rather than longitudinal, design. SIRIUS courses were implemented over a three-year period, with introductory courses introduced first and advanced courses introduced last, providing more time to collect data in the early-implementation courses. All advanced courses, with the exception of one, ran only once during the evaluation time frame. This was in part due to wildfires, which closed the campus during one semester that surveys were being administered to advanced students. Advanced course section sizes were small to begin with, and at the time, these courses were run only once per year, which further limited data collection. However, the advanced-level dataset included student responses from four to five distinct classes taught by different instructors, which may have served as a strength to the study design (i.e., more repeated treatments). Limited numbers, however, prevented us from addressing other factors that could have influenced student's perceptions, such as the number and combination of SIRIUS CURE courses completed. Our data do suggest that each level is composed of students that were similar demographically, and that these CUREs are impacting students with different identities in a similar manner; however, limited numbers across all sections prevented a thorough exploration of demographic variables that could have influenced perceptions.

Another limitation is the lack of uniformity in training for the faculty involved in teaching the SIRIUS courses (Table 2). Differences in instructors' degree of training and level of participation in CURE designs could affect students' perceptions of CUREs. Thus, observed differences in student perceptions across course levels may partially stem from instructor training and "buy-in." Specifically, the high LCAS scores and sophisticated responses to the "real science" question in advanced courses could be explained partly by the courses being taught solely by SIRIUS FLC participants, all of whom were tenure-track faculty. For intermediate and introductory courses, some lecturers and GTAs participated in SIRIUS professional development and course-design activities. However, introductory courses had a greater percentage of lecturers and GTAs who did not participate in the faculty learning community, and thus may have presented and facilitated the CUREs differently. That said, all course coordinators for multi-section courses were part of the FLC and performed some training for the lecturers and GTAs teaching the CUREs.

Finally, we reiterate that differences observed between levels cannot be attributed to a single factor. Characteristics of the CUREs, such as length of time (Table 2) and degree of autonomy differed between levels. Students' accumulated knowledge and skills (including experience gained in other SIRIUS courses) differed between levels, as well. Last, class sizes and the degree of connection between lecture and lab sections differed between courses. While we are unable to pinpoint exact causes for differences, we are nonetheless encouraged by students' perception of authenticity, especially in the advanced levels.

Future Directions

This study was largely motivated by our aim to assess the SIRIUS curricula designed by faculty in our multi-year learning community to help inform a second phase of the project. In the newly-funded SIRIUS II project, we are expanding our model to 10 STEM disciplines and five institutions – one four-year institution and four community colleges – united around a shared

vision to train students with authentic and interdisciplinary experiences. In the assessment of SIRIUS II, we have the opportunity to explore questions that arose from this work, such as the relationship between the perceived authenticity and student outcomes or benefits. As the SIRIUS project was initiated to solve a problem related to limited research opportunities and inequitable access to authentic experiences, we also seek to determine whether the SIRIUS program is shifting the demographics of students who perform extracurricular research to include more students from historically and currently marginalized populations.

CONCLUSION

Our study indicates that students progressing through a multi-level CURE program recognize the authenticity of their laboratory work and perceive engaging in activities related to Collaboration, Discovery/Relevance, and Iteration. They also perceive an increased degree of independence and more opportunities to fail as they progress from introductory to advanced coursework. With this report, we add to the literature on CUREs, and highlight the potential benefits of a scaffolded CURE curriculum. We also confirm previous findings that integrating autonomy and opportunities for failure are especially salient elements that influence students' perceptions about science/research authenticity. Lastly, we highlight the benefits of using a triangulation mixed methods design, with different instruments addressing similar questions, to support findings and understand nuanced data. While our methods provided us with useful data about student perceptions of their CUREs, future work is needed to understand the influence of positive (or negative) student perceptions of authenticity on other outcomes. The question remains, do students who perceive their lab experiences as authentic science benefit more from CUREs than students who do not?

ACKNOWLEDGMENTS

The authors thank Dr. Beth Schussler and Dr. Maryrose Weatheron for their thoughtful comments and feedback on versions of this paper. We thank the editor and the reviewers for their helpful feedback throughout the revision process. We also thank the National Science Foundation and the William M. Keck foundation for funding the SIRIUS Project.

REFERENCES

- Allen, W. E., Hosbein, K. N., Kennedy, A. M., Whiting, B., & Walker, J. P. (2021). Design and implementation of an organic to analytical CURE sequence. *Journal of Chemical Education*, 98(7), 2199–2208. <https://doi.org/10.1021/acs.jchemed.1c00129>
- Asai, D. (2020). Race Matters. *Cell*, 181(4), 754–757. <https://doi.org/10.1016/j.cell.2020.03.044>
- Auchincloss, L. C., Laursen, S. L., Branchaw, J. L., Eagan, K., Graham, M., Hanauer, D. I., ... & 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>
- 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.14-06-0099>
- Beck, C., Butler, A., & Burke da Silva, K. (2014). Promoting Inquiry-Based Teaching in Laboratory Courses: Are We Meeting the Grade? *CBE—Life Sciences Education*, 13(3), 444–452. <https://doi.org/10.1187/cbe.13-12-0245>

- Beck, C. W., Cole, M. F., & Gerardo, N. M. (2023). Can We Quantify If It's a CURE? *Journal of Microbiology & Biology Education*, 24(1), 1–6. <https://doi.org/10.1128/jmbe.00210-22>
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107(2), 238–246. <https://doi.org/10.1037/0033-2909.107.2.238>
- Brownell, S. E., Hekmat-Scafe, D. S., Singla, V., Chandler Seawell, P., Conklin Imam, J. F., Eddy, S. L., ... & Cyert, M. S. (2015). A high-enrollment course-based undergraduate research experience improves student conceptions of scientific thinking and ability to interpret data. *CBE—Life Sciences Education*, 14(2), ar21. <https://doi.org/10.1187/cbe.14-05-0092>
- Brownell, S. E., Kloser, M. J., Fukami, T., & Shavelson, R. (2012). Undergraduate biology lab courses: comparing the impact of traditionally based “cookbook” and authentic research-based courses on student lab experiences. *College Science Teaching*, 41(4) 36–45.
- Buchanan, A. J., & Fisher, G. R. (2022). Current status and implementation of science practices in course-based undergraduate research experiences (CUREs): a systematic literature review. *CBE—Life Sciences Education*, 21(4), ar83. <https://doi.org/10.1187/cbe.22-040069>
- Buck, L., Bretz, S., & Towns, M. (2008). Characterizing the level of inquiry in the undergraduate laboratory. *Journal of College Science Teaching*, 37(7), 52–58.
- Byrne, B. M. (1994). *Structural equation modelling with EQS and EQS/Windows: basic concepts, applications, and programming*. Thousand Oaks, CA: Sage.
- Carlone, H. B., & Johnson, A. (2007). Understanding the science experiences of successful women of color: Science identity as an analytic lens. *Journal of Research in Science Teaching*, 44(8), 1187–1218. <https://doi.org/10.1002/tea.20237>
- Chen, X., & Ho, P. (2012, October 10). Stem in postsecondary education: Entrance, attrition, and coursetaking among 2003–2004 beginning post-secondary students. *National Center for Education Statistics*, Retrieved February 5, 2021, from <https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2013152>
- Cooper, K. M., Blattman, J. N., Hendrix, T., & Brownell, S. E. (2019). The impact of broadly relevant novel discoveries on student project ownership in a traditional lab course turned CURE. *CBE—Life Sciences Education*, 18(4), ar57. <https://doi.org/10.1187/cbe.19-06-0113>
- Cooper, K. M., Cala, J. M., & Brownell, S. E. (2021). Cultural capital in undergraduate research: An exploration of how biology students operationalize knowledge to access research experiences at a large, public research-intensive institution. *International Journal of STEM Education*, 8(1), 6. <https://doi.org/10.1186/s40594-020-00265-w>
- Corwin, L. A., Graham, M. J., & Dolan, E. L. (2015). Modeling course-based undergraduate research experiences: an agenda for future research and evaluation. *CBE—Life Sciences Education*, 14(1), es1. <https://doi.org/10.1187/cbe.14-10-0167>
- Corwin, L. A., Runyon, C. R., Ghanem, E., Sandy, M., Clark, G., Palmer, G. C., ... & Dolan, E. L. (2018). Effects of discovery, iteration, and collaboration in laboratory courses on undergraduates' research career intentions fully mediated by student ownership. *CBE—Life Sciences Education*, 17(2), ar20. <https://doi.org/10.1187/cbe.17-07-0141>
- Corwin, L. A., Runyon, C., Robinson, A., & Dolan, E. L. (2015). The laboratory course assessment survey: A tool to measure three dimensions of research-course design. *CBE—Life Sciences Education*, 14(4), ar37. <https://doi.org/10.1187/cbe.15-03-0073>
- Cruz, C. L., Holmberg-Douglas, N., Onuska, N. P. R., McManus, J. B., MacKenzie, I. A., Hutson, B. L., ... & Nicewicz, D. A. (2020). Development of a large-enrollment course-based research experience in an undergraduate organic chemistry laboratory: Structure–function relationships in pyrylium photoredox catalysts. *Journal of Chemical Education*, 97(6), 1572–1578. <https://doi.org/10.1021/acs.jchemed.9b00786>
- CUREnet. CUREnet. Retrieved July 3, 2020, from <https://serc.carleton.edu/curennet/index.html>
- Dedoose (2021). *Dedoose Version 9.0.17, cloud application for managing, analyzing, and presenting qualitative and mixed method research data*. Los Angeles, CA: SocioCultural Research Consultants, LLC. www.dedoose.com
- Dinno, A. (2017). *_dunn.test: Dunn's test of multiple comparisons using rank sums_R package version 1.3.5*, Retrieved May 15, 2021, from <https://CRAN.R-project.org/package=dunn.test>
- Dinno, A. (2015). Nonparametric pairwise multiple comparisons in independent groups using Dunn's test. *The Stata Journal: Promoting Communications on Statistics and Stata*, 15(1), 292–300. <https://doi.org/10.1177/1536867X1501500117>
- Dolan, E., & Weaver, G. (2021). *A Guide to Course-based Undergraduate Research Developing and Implementing CUREs in the Natural Sciences*. New York, NY: Macmillan Learning.
- Eagan, M. K., Hurtado, S., Chang, M. J., Garcia, G. A., Herrera, F. A., & Garibay, J. C. (2013). Making a difference in science education: the impact of undergraduate research programs. *American Educational Research Journal*, 50(4), 683–713. <https://doi.org/10.3102/0002831213482038>
- Easterby-Smith, M., Thorpe, R., & Jackson, P. R. (2012). *Management research* (4th ed.). Thousand Oaks, CA: SAGE.
- Fajardo, H. L. (2022). *Examining student perceptions and curricular outcomes of a series of course-based undergraduate research experiences (Unpublished master's thesis)*. Sacramento, California: California State University.
- Fan, X., Thompson, B., & Wang, L. (1999). Effects of sample size, estimation methods, and model specification on structural equation modeling fit indexes. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 56–83. <https://doi.org/10.1080/10705519909540119>
- Gin, L. E., Rowland, A. A., Steinwand, B., Bruno, J., & Corwin, L. A. (2018). Students who fail to achieve predefined research goals may still experience many positive outcomes as a result of CURE participation. *CBE—Life Sciences Education*, 17(4), ar57. <https://doi.org/10.1187/cbe.18-03-0036>
- Goodwin, E. C., Anokhin, V., Gray, M. J., Zajic, D. E., Podrabsky, J. E., & Shortlidge, E. E. (2021). Is this science? Students' experiences of failure make a research-based course feel authentic. *CBE—Life Sciences Education*, 20(1), ar10. <https://doi.org/10.1187/cbe.20-07-0149>
- Hanauer, D. I., Graham, M. J., & Hatfull, G. F. (2016). A measure of college student persistence in the sciences (PITS). *CBE—Life Sciences Education*, 15(4), ar54. <https://doi.org/10.1187/cbe.15-09-0185>
- 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. <https://doi.org/10.1002/sce.20173>
- Jonassen, D. H., & Rohrer-Murphy, L. (1999). Activity theory as a framework for designing constructivist learning environments. *Educational Technology Research and Development*, 47(1), 61–79. <https://doi.org/10.1007/BF02299477>
- Jones, M. T., Barlow, A. E. L., & Villarejo, M. (2010). Importance of undergraduate research for minority persistence and achievement in biology. *The Journal of Higher Education*, 81(1), 82–115. <https://doi.org/10.1353/jhe.0.0082>
- Kearney, M. (2022). Effect of changes in course delivery format in response to the COVID-19 pandemic on self-assessed learning gains in a large introductory biology CURE. *Journal of STEM Education: Innovations and Research*, 23(2), 35–44. <https://www.jstem.org/jstem/index.php/JSTEM/article/view/2572>
- Kuh, G. (2008). *High-impact educational practices: what they are, who has access to them, and why they matter*. Washington, DC: Association of American Colleges & Universities.
- Lansverk, A. L., Lichti, D., Blinka, K. W., & Callis-Duehl, K. L. (2020). Comparing the outcomes of “Pre-CURE” compared to inquiry-based introductory biology labs. *Bioscene: The Journal of College Biology Teaching*, 46(1), 10–22.
- Lave, J., & Wenger, E. (1991). *Situated learning: legitimate peripheral participation*. Cambridge, UK: Cambridge University Press.
- Linn, M. C., Palmer, E., Baranger, A., Gerard, E., & Stone, E. (2015). Undergraduate research experiences: impacts and opportunities. *Science*, 347(6222), 1261757. <https://doi.org/10.1126/science.1261757>
- Lopatto, D. (2007). Undergraduate research experiences support science career decisions and active learning. *CBE—life sciences education*, 6(4), 297–306. <https://doi.org/10.1187/cbe.07-06-0039>
- Lopatto, D., Alvarez, C., Barnard, D., Chandrasekaran, C., Chung, H.-M., Du, C., ... & Elgin, S. C. R. (2008). Genomics education partnership. *Science*, 322(5902), 684–685. <https://doi.org/10.1126/science.1165351>
- Lopatto, D., Rosenwald, A. G., DiAngelo, J. R., Hark, A. T., Skerritt, M., Wawersik, M., ... & Elgin, S. C. R. (2020). Facilitating growth through frustration: using genomics research in a course-based undergraduate

- research experience. *Journal of Microbiology & Biology Education*, 21(1), 40. <https://doi.org/10.1128/jmbe.v21i1.2005>
- Martin, A., Rechs, A., Landerholm, T., & McDonald, K. (2021). Course-based undergraduate research experiences spanning two semesters of biology impact student self-efficacy but not future goals. *Research and Teaching*, 50(4) 33–47.
- McDonald, K. K., & Landerholm, T. E. (2018). Making researchers of STEM undergraduates. *Scientia, STEM Education*, Retrieved April 14, 2021, from <http://www.scientia.global/dr-kelly-mcdonald-dr-thomaslanderholm-making-re-searchers-of-stem-undergraduates>
- McDonald, K., Martin, A. R., Watters, C. P., & Landerholm, T. (2019). A faculty development model for transforming a department's laboratory curriculum with course-based undergraduate research experiences. *Journal of College Science Teaching*, 48(3), 14–23. https://doi.org/10.2505/4/jcst19_048_03_14
- Microsoft Corporation. (2018). Microsoft Excel. Retrieved August 1, 2021, from <https://office.microsoft.com/excel>
- Munawar, F. (2015). Assessment of awareness, perceptions, skills and motivation of undergraduate medical students 'research in physiology-a pilot study. *MOJ Anatomy & Physiology*, 1(2), 00008.
- Munn, M., Knuth, R., Van Horne, K., Shouse, A. W., & Levias, S. (2017). How do you like your science, wet or dry? How two lab experiences influence student understanding of science concepts and perceptions of authentic scientific practice. *CBE—Life Sciences Education*, 16(2), ar39. <https://doi.org/10.1187/cbe.16-04-0158>
- Nagda, B., Gregerman, S., Lerner, J., von Hippel, W., & Jonides, J. (1998). Undergraduate student-faculty research partnerships affect student retention. *The Review of Higher Education*, 22(1), 55–72. <https://doi.org/10.1353/rhe.1998.0016>
- Nunnally, J. C. (1978). *Psychometric theory*. New York, NY: McGraw-Hill.
- Olimpo, J. T., Fisher, G. R., & DeChenne-Peters, S. E. (2016). Development and evaluation of the Tigriopus course-based undergraduate research experience: impacts on students' content knowledge, attitudes, and motivation in a majors introductory biology course. *CBE—Life Sciences Education*, 15(4), ar72. <https://doi.org/10.1187/cbe.15-11-0228>
- R Core Team (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Retrieved February 5, 2023, from <https://www.R-project.org/>
- Rahm, J., Miller, H. C., Hartley, L., & Moore, J. C. (2003). The value of an emergent notion of authenticity: examples from two student/teacher-scientist partnership programs. *Journal of Research in Science Teaching*, 40(8), 737–756. <https://doi.org/10.1002/tea.10109>
- Revelle, W. (2022). *psych: Procedures for psychological, psychometric, and personality research*. Evanston, Illinois: Northwestern University.
- R package version 2.2.9, Retrieved August 11, 2022, from <https://CRAN.R-project.org/package=psych>
- Rizopoulos, D. (2006). "ltn: An R package for latent variable modelling and item response theory analyses." *Journal of Statistical Software*, 17(5), 1–25.
- 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), ar20. <https://doi.org/10.1187/cbe.16-03-0117>
- Rossee, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1–36.
- Rowland, S., Pedwell, R., Lawrie, G., Lovie-Toon, J., & Hung, Y. (2016). Do we need to design course-based undergraduate research experiences for authenticity? *CBE—Life Sciences Education*, 15(4), ar79. <https://doi.org/10.1187/cbe.16-02-0102>
- Saldana, J. (2021). "Quantitizing" the qualitative. In *The coding manual for qualitative researchers* (4th ed., pp. 414). Thousand Oaks, CA: SAGE Publications Ltd.
- Sathy, V., Nasiri, M., Strauss, C., & Hutson, B. (2020). The CURE for broadening participation in undergraduate teaching. In Ober, T. M., Che, E., Brodsky, J. E., Raffaele, C., & Brooks, P. J. (Eds.), *How we teach now: The GSTA guide to transformative teaching* (pp. 429–444). Society for the Teaching of Psychology. <http://teachpsych.org/ebooks/howweteachnow-transformative>
- 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(4), 493–534. <https://doi.org/10.1002/sce.10131>
- Spell, R. M., Guinan, J. A., Miller, K. R., & Beck, C. W. (2014). Redefining authentic research experiences in introductory biology laboratories and barriers to their implementation. *CBE—Life Sciences Education*, 13(1), 102–110. <https://doi.org/10.1187/cbe.13-08-0169>
- Weaver, G. C., Russell, C. B., & Wink, D. J. (2008). Inquiry-based and research-based laboratory pedagogies in undergraduate science. *Nature Chemical Biology*, 4(10), 577–580. <https://doi.org/10.1038/nchembio1008-577>
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. New York, NY: Springer-Verlag. Retrieved December 10, 2022, from <https://ggplot2.tidyverse.org>
- Wiggins, B. L., Sefi-Cyr, H., Lily, L. S., & Dahlberg, C. L. (2021). Repetition is important to students and their understanding during laboratory courses that include research. *Journal of Microbiology & Biology Education*, 22(2), e00158–21. <https://doi.org/10.1128/jmbe.00158-21>