3-1-2021

Is This Science? Students' Experiences of Failure Make a Research-Based Course Feel Authentic

Emma C. Goodwin
Portland State University, e.c.goodwin@pdx.edu

Vladimir Anokhin
Portland State University

MacKenzie J. Gray
Portland State University, mackgray@pdx.edu

Daniel E. Zajic
Portland State University, zajicde@gmail.com

Jason Podrabsky
Portland State University, podrabsj@pdx.edu

See next page for additional authors

Citation Details

Goodwin, E. C., Anokhin, V., Gray, M. K. 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

Authors

Emma C. Goodwin, Vladimir Anokhin, MacKenzie J. Gray, Daniel E. Zajic, Jason Podrabsky, and Erin E. Shortlidge

Is This Science? Students' Experiences of Failure Make a Research-Based Course Feel Authentic

Emma C. Goodwin, Vladimir Anokhin, MacKenzie J. Gray, Daniel E. Zajic, Jason E. Podrabsky, and Erin E. Shortlidge*

Biology Department, Portland State University, Portland, OR 97201

ABSTRACT

Course-based undergraduate research experiences (CUREs) and inquiry-based curricula both expose students to the scientific process. CUREs additionally engage students in novel and scientifically relevant research, with the intention of providing an "authentic" research experience. However, we have little understanding of which course design elements impact students' beliefs that they are experiencing "authentic" research. We designed a study to explore introductory biology students' perceptions of research authenticity in CURE and inquiry classes. Using the Laboratory Course Assessment Survey, we found that students in CURE sections perceived higher levels of authentic research elements than students in inquiry-based sections. To identify specific factors that impact perceptions of research authenticity, we administered weekly reflection questions to CURE students. Coding of reflection responses revealed that experiences of failure, iteration, using scientific practices, and the relevant discoveries in their projects enhanced students' perceived authenticity of their research experiences. Although failure and iteration can occur in both CUREs and inquiry-based curricula, our findings indicate these experiences—in conjunction with the Relevant Discovery element of a CURE—may be particularly powerful in enhancing student perceptions of research authenticity in a CURE.

INTRODUCTION

Undergraduate research experiences have the potential to increase student motivation, interest, and retention in science, technology, engineering, and mathematics (STEM) fields—particularly for students who are traditionally underrepresented in the sciences (Laursen *et al.*, 2010; Eagan *et al.*, 2013; National Academies of Sciences, Engineering, and Medicine [NASEM], 2015). Universities have therefore been tasked with increasing opportunities for STEM students to participate in these often-transformative research experiences (American Association for the Advancement of Science, 2011; President's Council of Advisors on Science and Technology, 2012; Bangera and Brownell, 2014). However, many students do not have the option or ability to participate in traditional research apprenticeships due to various constraints (Bangera and Brownell, 2014), leading to increasing efforts to integrate discovery-based courses into the curricula (NASEM, 2015). Such courses are thought to be particularly impactful for students at the introductory level—the point at which many students leave the STEM degree path (Graham *et al.*, 2013).

Intentionally engaging students in their own learning can positively impact student outcomes such as exam performance and student buy-in (Freeman *et al.*, 2014; Cavanagh *et al.*, 2016). Buy-in can manifest both in endorsement and in attitudes toward active learning and has been linked to increased engagement and improved course performance (Cavanagh *et al.*, 2016). Further, student recognition that authentic research elements have been integrated into their courses can result in an increased interest and motivation by students to do research (Vereijken *et al.*, 2016, 2019). Thus,

student buy-in to the authenticity of a research experience may have the potential to increase engagement, motivation, and performance. One goal of developing discovery-based curricula should therefore be engaging students in a research experience that is authentic—from the perspectives of both educators and (potentially more importantly) students.

Designing research-based curricula raises the question: What should an authentic research experience in an undergraduate course look like? Research in the space of an undergraduate classroom may look inherently different from research performed by a research scientist, in that it is inevitably constrained by the structural elements of a course, such as class schedule, equipment availability, cost of course materials, and finite length of the academic term (Spell *et al.*, 2014; Bakshi *et al.*, 2016; Shortlidge *et al.*, 2016; Thompson *et al.*, 2016; Govindan *et al.*, 2020). These constraints necessitate redefining what "authentic" research looks like when adapted for the classroom. Previous research aimed to define research authenticity in the space of a science classroom from the perspectives of educators and education researchers (Spell *et al.*, 2014; Rowland *et al.*, 2016). Representatives of the Course-based Undergraduate Research Experiences Network (CURE.net) met in 2013 to create a defining framework for elements inherent to course-based undergraduate research experiences (Auchincloss *et al.*, 2014). However, efforts to date to define what authentic research practices look like have focused on the perspectives of experts, rather than the perceptions of students. It is unclear which (if any) design elements of courses facilitate students believing that what they are doing in their lab course is "authentic" research, and whether those perspectives align with a course designer's intended outcomes (Corwin *et al.*, 2015b). Unpacking the elements that allow students to buy into the authenticity of their lab courses will deepen our understanding of the elements that make research-based curricula a valuable experience for undergraduate STEM students.

Expert Perceptions of Authenticity: Is Science a Product or a Process?

While this study explores student perceptions of research authenticity in the classroom, we aim to frame our work within the diverse beliefs that educators hold regarding course design elements inherent to classroom-based scientific research. Rowland *et al.* (2016) compiled papers from the research literature in which authors (often STEM education researchers) provided their own definitions of what makes for "authentic science" in educational contexts. The authors analyzed 26 definitions of research authenticity and found that the top reported elements (according to the researchers) included: experiencing the process and practice of science (15 of 26 definitions), ownership/personal relevance to students (seven of 26), engaging students in experimental design (six of 26), and novel/publishable results and communication (both found in four of 26 definitions) (Rowland *et al.*, 2016).

As described in Rowland *et al.* (2016), some researchers suggest that there are two modes of thinking about authentic research in the classroom: 1) science as a "product" and 2) science as a "process." For example, in a national survey of introductory biology lab instructors, researchers found that faculty tend to gravitate to one of two distinct conceptions of authentic research in the classroom—one in which students have the goal of addressing novel questions and generating novel results (the "products" of science) or one in which students experience the process of science by participating in activities such as experimental design and data collection/analysis, without a goal of producing relevant scientific data (Spell *et al.*, 2014). A similar dichotomy is proposed by Barab and Hay (2001), who suggest that authentic research experiences can be either "participatory," in which students actually participate in an expert scientist's research program and assist in the production of research (working on "products" of science); or "simulated," in which students conduct scientific activities and thereby have the opportunity to simulate being an expert scientist (practicing the "process" of science). There are clear parallels between these two models of authentic research with respect to inquiry and research-based courses in undergraduate biology laboratory classrooms (summarized in Table 1).

It is presumptuous to assume that undergraduates—especially those new to research—and experts hold the same beliefs about research authenticity. For example, a multi-institutional study of 665 students and their instructors in 39 different inquiry lab courses found little relation between student and instructor perceptions of what happens in the lab classroom (Beck and Blumer, 2016). Further, it is unlikely that there is a singular context that students will uniformly perceive as "authentic"—Rahm and colleagues argue that the perception of authenticity can "emerge" for different students in different educational contexts (Rahm *et al.*, 2003). It is therefore critical to explore student perceptions of research authenticity in multiple educational contexts where research experiences are fostered, and here we consider both inquiry-based curricula and course-based undergraduate research experiences (CUREs).

TABLE 1. Alignment of Inquiry and CURE models with existing frameworks of authentic research in the science lab

Authenticity framework	Inquiry	CUREs
Authenticity can be simulated or participatory (Barab and Hay, 2001)	Students **simulate** the activities of an expert researcher.	Students **participate** in an expert's research project.
Authentic research includes the process or products (novel questions/results) of science (Spell *et al.*, 2014)	Prioritizes that students experience the **process of science** over answering novel questions.	Prioritizes that students seek to generate **novel results (products of science)** over experiencing the process of science.
CURE research dimensions (Auchincloss *et al.*, 2014)	Students may engage in **Scientific Practices, Collaboration,** and **Iteration**.	Students engage in novel **Relevant Discovery** in addition to Scientific Practices, Collaboration, and Iteration.

Bringing Authentic Research Elements into the Classroom

Inquiry-Based Courses. The last three decades have seen a large shift in undergraduate biology lab courses replacing cookbook-style labs with discovery-based courses that incorporate elements of inquiry and research into the classroom (Hofstein and Lunetta, 2004; Sundberg *et al.*, 2005; NASEM, 2015). In cookbook labs, students engage in "confirmatory" activities, in which all necessary information is provided to students, there is a "correct" outcome for students, and/or the students are learning a lab technique and essentially following a recipe (Domin, 1999; Hofstein and Lunetta, 2004; Buck *et al.*, 2008). In contrast, inquiry engages students in activities that allow them to develop their own scientific knowledge and understanding of the process of science through participation in many of the activities that research scientists regularly practice (National Research Council, 1996; Domin, 1999). The label "inquiry" applies to a broad range of course structures and design elements in the context of an undergraduate biology classroom, and there is no single agreed-upon definition of what an inquiry course looks like (Buck *et al.*, 2008). The relative control that students have over their activities in any inquiry course can vary greatly, from "structured" inquiry courses, in which students are guided through the majority of their work, to "open" courses, in which students have the autonomy to design their own research methods, collect and analyze data, and communicate their results (Buck *et al.*, 2008). Students in "authentic" inquiry courses may have the opportunity to develop their own research questions, though there is little expectation that students in these courses will produce publication-quality data or ask questions that are novel to the scientific community (Domin, 1999; Buck *et al.*, 2008; Spell *et al.*, 2014; Brownell and Kloser, 2015).

In Table 1, we outline our interpretation of how different discovery-based course designs align with the previously described models of authentic research proposed by Barab and Hay (2001) and Spell *et al.* (2014). When classifying inquiry courses within the context of Barab and Hay's (2001) simulated versus participatory authenticity framework, we believe that inquiry-style experiences offer students the chance to *simulate* the experiences of an expert scientist, because students are engaging in the process of science and often have some control over their study design and methods. Inquiry courses may therefore be "authentic," in the sense that students can engage in the same practices as an expert scientist (the "process" of science), even though students are not producing novel and/or relevant data (Spell *et al.*, 2014; Table 1).

CUREs. Increasingly prevalent in the literature is a focus on courses in which students *do* produce potentially publishable data (e.g., see Auchincloss *et al.*, 2014; Corwin *et al.*, 2015a; NASEM, 2015; Shortlidge *et al.*, 2016; Ballen *et al.*, 2017). Involving students in research through a CURE exposes students to the use of multiple Scientific Practices, Discovery, Broader Relevance, Collaboration, and Iteration (Auchincloss *et al.*, 2014). While students in an inquiry activity may engage in one or more of these practices, the opportunities for novel discoveries that have relevance to the scientific community specifically distinguish CUREs from inquiry courses (Auchincloss *et al.*, 2014; Brownell and Kloser, 2015; Cooper *et al.*, 2017). Recent work has suggested that Discovery and Broader Rele-

vance are difficult to disentangle in the context of a CURE (see Brownell and Kloser, 2015; Corwin *et al.*, 2015b; Cooper *et al.*, 2019). We follow the lead of Cooper *et al.* (2019) in considering these features as a single item: Broadly Relevant Novel Discovery, which we hereafter refer to as "Relevant Discovery."

Like inquiry-based curricula, CUREs vary greatly in design, but generally fall into one of two categories: 1) independent CUREs, often designed by researchers and/or instructors and frequently based around their research program/interests, which can result in locally or broadly relevant data; or 2) large-scale "network" CUREs, designed for instructors to implement with relative ease (Shortlidge *et al.*, 2017). In both models, students are producing potentially publishable research, though they may have varied control over their research questions and methodological choices (Brownell and Kloser, 2015). Spell and colleagues (2014) cite several examples of independent and network CUREs that emphasize the "science as a product" model of authentic research, in which the aim of participating in the CURE is the production or analysis of relevant novel data, and many CURE instructors aim for this outcome (Shortlidge *et al.*, 2016). When the goal of a CURE is for students to contribute to a larger scientific effort, Barab and Hay's "participatory" rather than "simulated" model of authentic science is emphasized (Barab and Hay, 2001). Within the CURE framework, use of multiple Scientific Practices, Iteration, and Collaboration represent the "science as a process" model that students experience in simulated research experiences (Table 1). The combination of these "science as a process" elements with Relevant Discovery aligns with the model of authenticity that emphasizes the products of science (seeking to answer novel questions and generate relevant results)—a goal of the participatory research model.

Do Students Buy into the Authenticity of Their Classroom Lab Experiences?

The educational contexts in which student perceptions of authenticity can emerge could be quite different from what experts may perceive to be authentic research experiences (Rahm *et al.*, 2003). Indeed, students in both CURE and inquiry courses use the words "real," "actual," and "genuine" to describe their experiences (Rowland *et al.*, 2016), indicating that students may perceive their experiences to be authentic regardless of whether they are participating in scientific research or simulating the scientific process (Barab and Hay, 2001).

There is little research into the specific activities that promote undergraduate students' perceptions of participating in authentic research. A study of nearly 300 high school students who participated in either dry lab (using a database to explore questions about factors that could influence smoking habits) or wet lab (using molecular techniques to genotype DNA from human subjects) research found that students in the dry lab reported participating in a number of scientific activities at a significantly higher level than in the wet lab, including: coming up with their own research question, testing hypotheses, analyzing data, and drawing conclusions. In contrast, students in the wet lab only reported using the same tools and equipment as scientists do at a significantly higher rate than dry-lab students. Despite the many scientific activities that dry-lab students reported participating in compared with wet-lab students, students in the wet lab had a higher perception that their

experience was more similar to what "real scientists" do (Munn *et al.*, 2017). Therefore, simply using scientific tools and equipment—an important component of both inquiry and CURE courses—may be a critical factor impacting students' perceptions that they are participating in "authentic" science.

In this study, we compare how students in a CURE and an inquiry course (hereafter referred to as "CURE students" and "inquiry students," respectively) experience authentic research elements in their curricula, and we identify factors that influence CURE students' perceptions of research authenticity. We quantitatively compare CURE and inquiry students' perceptions of experiencing the different dimensions of research using the CURE framework, with the hypothesis that CURE students will perceive higher levels of Collaboration, Discovery, and Iteration. As CUREs are designed such that students experience both the "process" and "products" of science—both presumed dimensions of authenticity—we developed a series of open-response questions for CURE students to reflect on their course experiences and unpack what contributed to or detracted from the perception that the classroom research experience was authentic. We evaluate our findings of student perceptions of authentic research in relation to how authenticity is described by practitioners in the literature.

METHODS

Course Structure and Study Participants

We conducted this study at a large, urban, public university in the Pacific Northwest, with a largely nontraditional student population with students of various ages and prior college experiences. For this study, we worked with students in the third term of the 200-level introductory biology for majors laboratory sequence during the Spring 2018 academic term. This was a one-unit course associated with a large introductory biology lecture course, and labs were held for 3 hours per week throughout a 10-week quarter.

There were 21 lab sections led by graduate teaching assistants (GTAs). Students in all lab sections experienced the same conceptual and skill-building labs for the first 4 weeks of the term. In the remainder of the term, 17 of the lab sections continued with two more typical lab weeks, followed by a 4-week inquiry module. These "inquiry sections" were led by nine GTAs and involved 373 students. In the inquiry sections, students collaborated in small groups to design behavioral ecology experiments using sowbugs and had the autonomy to develop almost any experiment they wished to execute, given the available time and materials. Students were able to revise or repeat their experiments during the second week of the inquiry module. Students then conducted statistical analyses on both their team data and a larger data set collected from student groups across all inquiry lab sections, and groups designed PowerPoint presentations of their experiments and shared them with their lab sections at the end of the term. Students were not graded on the "success" of their experiments but rather on effort and their process of designing experiments and analyzing data to the best of their ability. We categorize this as an inquiry-based course, because students developed their own hypotheses and designed their own experiments, but their experiments were not necessarily novel and were not expected to produce potentially publishable data (Domin, 1999; Buck *et al.*, 2008; Spell *et al.*, 2014; Brownell and Kloser, 2015). Students therefore *simulated* the

process of science and experienced Collaboration and Iteration while using multiple Scientific Practices (Table 1).

Concurrently, four lab sections participated in a 6-week "killifish CURE" rather than the inquiry sequence. The killifish CURE lab sections were determined before enrollment opened for the term and were selected to allow for the CURE sections to run concurrently once a week in the afternoon and the evening to both minimize preparation and to allow the GTAs to assist one another. To control for instructor effect in the associated lecture course, we only allowed students enrolled in the larger daytime lecture section to enroll in the CURE sections, which was a minimal logistical barrier, as two of the CURE lab sections overlapped with the evening lecture. Because self-selection can impact student motivation (Rosenthal, 1965; Brownell *et al.*, 2013), we did not inform students during the enrollment period that certain sections would use the CURE curriculum. One week before the beginning of term, students in the CURE lab sections were informed that they were in a special lab section that would allow them to participate in research. Students were therefore able to switch lab sections if they desired. All but one student remained in their originally enrolled lab section. In this way, bias for self-selection into the CURE curriculum was minimized.

The CURE lab sections were led by two GTAs and involved 87 students. The killifish CURE was based on a biology faculty member's research program (J.E.P.) and was codeveloped with the instructor of record for the lecture and lab course, who is a biology faculty member and education researcher (E.E.S.). The CURE GTAs (E.C.G. and D.E.Z.) were advisees of the faculty leads and were closely involved with designing the CURE curriculum. In the killifish CURE, students designed two iterative rounds of experiments to test which biotic and abiotic factors can induce entrance into diapause (developmental arrest) in the embryos of *Austrofundulus limnaeus*, an annual killifish species that inhabits ephemeral ponds in Venezuela. CURE students participated in a brainstorming activity to develop novel hypotheses and experiments that would build on prior research on the topic, during which the GTAs subtly guided students toward a few predetermined experimental design options that course instructors believed could lead to potentially publishable data. Thus, the intention was for students to feel they had some autonomy in developing the research questions and experimental design, and the course instructors were able to ensure that student projects were feasible and could be accommodated at a large scale. Throughout the CURE, students collaborated in small groups, and as in the inquiry sections, students had the opportunity to revise, repeat, or expand on their experiments, and student grades were not impacted by the "success" of their experiments.

We designed the CURE to intentionally incorporate all CURE elements: Collaboration, Iteration, and use of multiple Scientific Practices, all in the context of Relevant Discovery (Auchincloss *et al.*, 2014). Our goal was for students to *participate* in faculty-driven research with the goal of *producing* novel and scientifically relevant data (Table 1). To scaffold Relevant Discovery into our curricula, we had students read a research paper from the faculty killifish researcher (J.E.P.), and we showed students a video and pictures highlighting research from the killifish lab to familiarize them with the research program they were contributing to. Both the faculty researcher

(J.E.P.) and the instructor of record (E.E.S.) visited the CURE sections, and students had the opportunity to directly discuss their projects with the faculty researcher and get feedback and advice on their experimental designs.

In scaffolding the CURE, we inherently introduced differences between the CURE and inquiry experiences that could impact direct comparisons between the course types, and we have made an effort to highlight these differences throughout this paper to increase transparency of the limitations of this study. For example, while both CURE and inquiry students were asked to do a similar amount of work in their respective labs, and all students worked in groups, CURE students were allowed to submit assignments that they completed as a group, while inquiry students completed their assignments individually. Because the CURE students needed separate lab periods to set up their experiments and collect their data, CURE students spent two more weeks on the CURE project compared with inquiry students, who could complete the entirety of their experiments (setup and data collection) within a single lab period. The nature of the assignments and assessments in the CURE sections were also slightly different, as they were designed to help students document and understand their experimental design and data collection, analysis, and interpretation. CURE students also answered weekly reflection questions (described later), which could have impacted their perceptions, as they prompted students to think about their course experiences.

All students enrolled in the labs were recruited to participate in a research study in the first week of the term, and in total 302 inquiry students (81% of total inquiry section enrollment) and 74 CURE students (85% of total CURE section enrollment) consented to be part of the research study. By consenting, students allowed researchers access to course assignments, surveys, institutional information, and their final lab and lecture grades. This study was approved by the university's Institutional Review Board (no. 184544).

Data Collection

We addressed our research questions with an embedded mixed-methods approach, in which we concurrently collected quantitative survey data from both CURE and inquiry students and written reflection responses from CURE students (Creswell, 2009). These data were collected to allow us to compare perceived levels of authentic research elements between the two course designs and to gain a deeper understanding of how students interpret research authenticity in a classroom setting.

Laboratory Course Assessment Survey. We used the Laboratory Course Assessment Survey (LCAS; Corwin *et al.*, 2015b), a 17-item instrument, to measure CURE and inquiry students' perceived levels of experiencing specific authentic research elements in their lab courses. The LCAS has previously been used to detect differences in student experiences across course types (Corwin *et al.*, 2015b, 2018; Cooper *et al.*, 2019; Esparza *et al.*, 2020) and was specifically designed to measure perceived participation in Collaboration, Discovery/Relevance (referred to here as "Relevant Discovery"), and Iteration activities. This allowed us to compare student perceptions of both "science as a process" (Collaboration and Iteration) and "science as a product" activities (Relevant Discovery). Students in the inquiry labs

were prompted to consider the sowbug experiments in answering the questions, while the CURE students were prompted to consider the killifish experiments. We predicted that CURE students would in general perceive higher levels of Collaboration, Relevant Discovery, and Iteration. We expected that one survey item (Relevant Discovery item 3: "I was expected to formulate my own research questions or hypothesis to guide an investigation") would behave inconsistently with our prediction, because CURE students were guided toward testing research questions that could feasibly lead to novel and potentially publishable data, whereas inquiry students were given carte blanche in forming hypotheses related to sowbug behavior.

The original survey was designed for students on the semester system, but because we are on a quarter system, we modified the response-scale options used for the Collaboration items to align with a more condensed course schedule. For example, the response option "Monthly" became "A couple of times, but not every lab period." The final version of our survey (Supplemental Material, Appendix 1) was reviewed by several undergraduate representatives of our student population and by GTAs of both the CURE and inquiry sections. We disseminated the survey online via Qualtrics to all lab students in the introductory biology course during the last week of the term, and students were offered 2 points of extra credit for taking the survey. In total, 201 inquiry students (67% of inquiry student participants) and 45 CURE students (61% of CURE student participants) responded to the survey.

CURE Student Reflections. To explore students' beliefs and feelings about participating in the CURE, we assigned students one to three weekly reflection questions as part of their regular quizzes throughout the 6-week CURE module. In total, 12 reflection questions were administered to students, and responses were graded by GTAs for completion rather than content. Because we were primarily interested in students' perceptions of research authenticity after they had experience with the CURE, we focused our analysis on nine questions that were administered in the final 3 weeks of the CURE (Table 2).

Data Analysis

LCAS Confirmatory Factor Analyses and t Tests. We administered the LCAS to CURE and inquiry lab students to measure perceptions of Collaboration, Relevant Discovery, and Iteration. Although the LCAS was developed and shown to produce valid data at other institutions for use with undergraduate STEM students, different student populations may interpret survey items in unique ways, and even minor modifications to any instrument could impact student responses (Barbera and VandenPlas, 2011). We therefore used confirmatory factor analysis (CFA) to collect evidence of construct validity by testing whether the latent construct structure of the instrument functions for our institutional population and course context (Hancock *et al.*, 2018). We specifically tested a correlated three-factor model with Collaboration, Relevant Discovery, and Iteration as separate latent factors (see Supplemental Material, Appendix 2). We used a robust maximum likelihood estimator with the Satorra-Bentler correction in all CFAs to correct for potential nonnormality in our item responses. While the maximum likelihood estimator assumes a continuous response scale, which is not ideal for data with fewer than five response

TABLE 2. CURE student reflection questions

CURE context	Question ID	Question text
Week 4 Students completed data collection and analysis from experiment 1 and monitored progress of experiment 2.	1	Last week the researcher who directs our CURE project stopped by to check in on your experimental progress. Were your interactions valuable? Why or why not?
	2	Last week our embryos did not develop as quickly as we were expecting and many unexpectedly died. How do you feel about the fact that we had to make last-minute changes to our experimental plan?
	3	What has been the most challenging aspect of this course so far for you?
Week 5 Students completed data collection and analysis from experiment 2.	4	Do you feel that you conducted real scientific research in lab this term? Why or why not?
	5	Do you see yourself as a scientist and/or a person who utilizes scientific principles and practices in your daily life? Please explain why/why not.
	6	Have your perceptions of what it means to do scientific research changed due to participating in the CURE portion of this lab course? If so, what has changed?
Week 6 Students presented their CURE projects to class.	7	If you had the opportunity to spend five more weeks in this lab, what would you want to do or learn with the extended time?
	8	Until this CURE, most of your previous introductory biology lab experiences involved lab activities that did not extend beyond a single lab period. Were there any difficulties or frustrations you faced due to the multiweek structure of the CURE lab project? Which format do you prefer?
	9	What skills that you practiced in this course were new to you? Describe the most useful skill you learned from this course, and why it is valuable to you.

categories and therefore likely underestimates our model fit (Hancock *et al.*, 2018), we chose to proceed with this estimator to maintain continuity with prior studies (e.g., Corwin *et al.*, 2018).

To determine the appropriate statistic to use as an estimate of the internal consistency of our instrument scales, we ran single-factor CFAs for each of the three factors using both a congeneric model (i.e., unrestricted factor loadings) and a tau-equivalent model (i.e., all factor loadings are forced to be equivalent; Komperda *et al.*, 2018b). The omega reliability coefficient is equivalent to Cronbach's alpha when factor loadings are equivalent but avoids bias introduced by Cronbach's alpha when factor loadings are independent (Komperda *et al.*, 2018a, b). We therefore report Cronbach's alpha as an estimate for reliability when the data–model fit met our study cutoffs (Confirmatory Fit Index [CFI] and Tucker-Lewis Index [TLI] ≥ 0.950, and root mean square error of approximation [RMSEA] ≤ 0.05; as suggested by Hancock *et al.*, 2018) under tau-equivalent conditions, and omega total when model fit met the study cutoffs only for the congeneric model.

Item scores for each construct were summed, and we used *t* tests to test for differences between sum construct scores for inquiry and CURE students and Hedge's *g* to calculate effect size. We also tested for differences between inquiry and CURE students in demographics and lab/lecture grades using chi-square tests of independence for categorical data and *t* tests for continuous data. Welch's *t* test was used whenever Bartlett's test for homogeneity of variance indicated that sample variances were unequal. All statistical analyses were conducted in R v. 3.6.2, using the base, lavaan, and userfriendlyscience packages (Rosseel, 2012; Peters, 2018; RStudio Team, 2019).

Qualitative Data Analysis of CURE Reflection Responses. Three researchers (E.C.G., V.A., M.J.G.) reviewed all CURE reflection responses and together established a coding scheme to capture the reoccurring sentiments in the responses. We developed the coding scheme using both a priori codes based on the CURE framework (Collaboration, Relevant Discovery, Iteration, and Scientific Practices; Auchincloss *et al.*, 2014) and initial structural coding, for which we created codes to describe ideas that were arising from the text responses (Saldana, 2015). Each code was a short label that encompassed a specific perception or experience that students described and was accompanied by a longer definition to clarify the code for the research team. For example, the code "Real Research: Iteration" was defined as: "Iteration, repeating experiments, or doing the experiment over a period of weeks contributes to student perceptions that the CURE was 'real research.'" The coding scheme was organized into thematically similar categories of codes (e.g., "Factors that contribute to perceptions that CURE is 'real research'"). While we developed codes that allowed for analysis of all written reflection responses, there were certain code categories that were only relevant to specific sets of questions. Within this work, we focus on code categories regarding students' perceptions about whether their CURE experiences felt like "real research." The three researchers coded all reflections independently in small sets and calculated percent agreement for each set. The final percent agreement for all coding data averaged between the three reviewers was 72%. Percent agreement calculations were used to ensure high coding standards were maintained among the team and to facilitate reflexive conversations throughout the coding process, rather than to formally quantify our reliability or divide labor between multiple coders (O'Connor and Joffe, 2020). All three researchers carefully discussed every code designation in all student reflections to consensus.

RESULTS

Demographics and Student Experiences

We collected institutional data for all study participants, and found that, on average, CURE students were slightly older than inquiry students (CURE mean age = 24.3 years, inquiry mean age = 22.8 years; Welch's $t = 2.023$, $df = 97.94$, $p = 0.05$). We did not detect any other significant demographic differences between the CURE and inquiry students (chi-square tests of independence, Supplemental Material, Appendix 3).

CURE and inquiry student lecture grades did not differ significantly from one another (CURE lecture grade average = 84.9%, inquiry lecture grade average = 86.3%; $t = 1.158$, $p = 0.25$). However, CURE students scored on average 2 percentage points more than inquiry students in the lab (CURE lab grade average = 96.3%, inquiry lab grade average = 94.3%, Welch's $t = 2.632$, $p < 0.01$). This is possibly due to the experiment-focused and collaborative CURE group assignments rather than the individual assignments expected from inquiry students.

CURE Students Perceive Higher Levels of Collaboration, Relevant Discovery, and Iteration

We collected descriptive statistics for each LCAS survey item to assess the normality of our data and found no items that displayed extreme deviations from normality (Supplemental Material, Appendix 4). We used a robust estimator in the CFAs to account for any moderate deviations from normality in our data. Single-factor CFAs indicated that omega total is an appropriate reliability statistic for all three scales, and all three scales had high internal consistency (Supplemental Material, Appendix 5). As predicted, within the single factor Relevant Discovery subscale, item 3 ("I was expected to formulate my own research questions or hypothesis to guide an investigation") had a substantially lower factor loading compared with the other Relevant Discovery items, and summary statistics (Supplemental Material, Appendix 4) indicate a reduced gap between CURE and inquiry students for this item. We discussed our theoretical concerns about this item with one of the LCAS authors and ultimately decided our theoretical and quantitative evidence was sufficient to omit this item from further data analysis with this study population. While the following analyses omit Relevant Discovery item 3, we found that presence or absence of this item has negligible effect on the three-factor model fit and the summed differences between CURE and inquiry students for the Relevant Discovery subscale.

We tested the a priori correlated three-factor model with Collaboration, Relevant Discovery, and Iteration as separate latent factors (see Supplemental Material, Appendix 2). Modification indices indicated a strong correlation between Iteration item 1 (I1) and the Relevant Discovery scale, indicating that I1 is not functioning as expected. We hypothesize that this could be due to I1's shared question stem with the Relevant Discovery items (Supplemental Material, Appendix 1). We therefore removed this item from the final analysis. Fit indices for the final model indicate that it was functioning appropriately for our student population (Table 3).

We summed the LCAS scores for each scale, using only the items included in our final model. While students in both CURE and inquiry lab sections perceived relatively high levels of Collaboration, Relevant Discovery, and Iteration, CURE students reported experiencing significantly higher levels of each

TABLE 3. Fit indices for LCAS CFA

Fit indices	Data–model fit	Accepted cutoff[a]
CFI	0.977	≥0.950
TLI	0.972	≥0.950
RMSEA (90% confidence)	0.047 (0.024–0.066)	≤0.050

[a] As suggested by Hancock et al. (2018).

construct in their laboratory course (t tests, $p < 0.001$; Figure 1; see Table 4 for test statistics). The largest effect size between inquiry and CURE students was seen for the Iteration scale, though there was also a medium effect size for the Relevant Discovery scale. In comparing these observed means for the LCAS factors between CURE and inquiry students, we ideally would have first conducted strict measurement invariance tests between the two groups to establish that error variances were similar across groups; however, our CURE student group was too small ($N = 45$) to conduct invariance tests (Rocabado et al., 2020).

CURE Students Perceive That Their Research Experience Is Authentic

We coded students' responses to the reflection question "Do you feel that you conducted real scientific research this term?" into three mutually exclusive categories. We found that the majority (76%) of CURE students believed that they conducted real scientific research and provided a variety of justifications for why their experience was "real," as exemplified by the following quote:
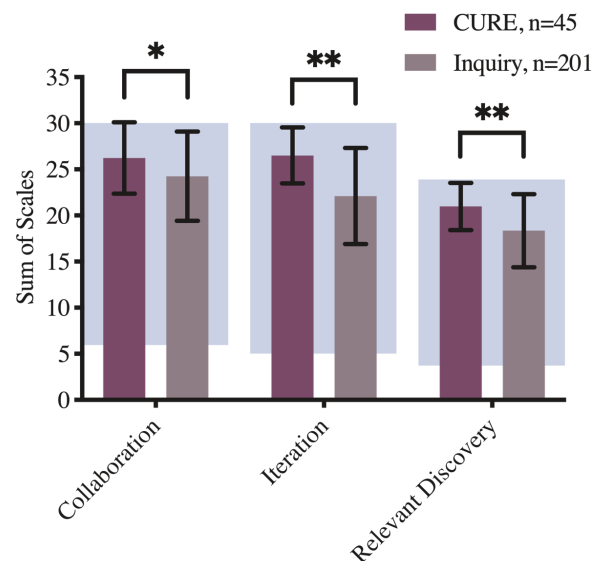

FIGURE 1. CURE students perceive significantly higher Collaboration, Relevant Discovery, and Iteration compared with their inquiry peers, as indicated by higher numbers for each scale (t tests: *$p < 0.05$; **$p < 0.001$; see Table 4 for test statistics). Background shading indicates potential score range of each summed scale: Relevant Discovery and Iteration were measured on a six-point Likert scale, while Collaboration was measured on a five-point Likert scale. Bars represent data mean ± SD.

TABLE 4. LCAS Collaboration, Iteration, and Relevant Discovery scores for CURE and inquiry students

Scale	Score range	Inquiry students *n* = 201		CURE students *n* = 45		Welch's *df*[a]	*t*	*p*	Hedge's *g*[b]
		Mean	SD	Mean	SD				
Collaboration	6–30	24.26	4.83	26.24	3.87	NA	2.58	0.011	0.42
Iteration	5–30	22.10	5.21	26.51	3.04	110.31	7.56	<0.001	0.90
Relevant Discovery	4–24	18.36	3.97	20.98	2.55	98.02	5.55	<0.001	0.70

[a]Welch's degrees of freedom were only used when the assumption of homogeneity of variance between inquiry and CURE students was not met (Bartlett's test) (Dalgaard, 2008).
[b]Effect size reference values are arbitrary, but in general a small effect size is below 0.5, a medium effect size is between 0.5 and 0.8, and a large effect size is greater than 0.8 (Hancock *et al.*, 2018).

> Yes, we did conduct real research. We went into these experiments not knowing what the outcome would be. We also got to design our own experiments. Some of them did not work, but that is how real research goes.

In total, 18% of CURE students were unsure whether they had conducted real scientific research and often provided thoughtful responses describing limitations they experienced during the course such as:

> Maybe we conducted real research. I feel that the sample size in our experiment is too small to be significant.

Only 7% of students reported that they had not experienced real research in the CURE:

> No, I feel like this CURE is too short for real scientific research. It is in a very controlled setting so in a way it does not feel real.

Several Factors Enhance the Perceived Authenticity of the Student Research Experience

To understand why students reported their research experiences did or did not feel real, we coded reflection responses to nine questions administered to students during the last 3 weeks of the CURE for justifications of student perceptions of research authenticity. On average, students described 1.9 unique reasons (SD = 1.2) justifying why they felt their experiences were authentic (summarized in Table 5). Unexpectedly, we found that experiences of Failure were the most cited explanation students provided for why the CURE felt like real research, which was discussed by 59% of CURE students. We refer to "Failure" as experiences wherein students are unable to successfully carry out a task to achieve a specific goal (Henry *et al.*, 2019). Students in the CURE all experienced failure during the term, as the majority of the killifish embryos they were working with perished, and very few student teams finished the term with interpretable results. Students were not graded on their experimental success and were able to repeat their experiments to try to achieve clearer results. These students rarely seemed discouraged by their experiences of failure, and sometimes even found them invigorating, as expressed in the following quote:

> I love that the experiment did not go as planned—I mean, sure, it is not ideal that a bunch of embryos died, however, this is how real science works. I am usually so bored in the assigned labs... [they] are carefully designed so that students get the "right" answer [in response to question 2; see Table 2 for question list].

Although students reflected on their experiences of failure unprompted, we also specifically prompted students in one question to discuss their feelings about the embryo die-off, which could have led to artificially inflated proportions of students using failure as a rationale for why their experience felt real. However, Failure clearly resonated with students as they considered the authenticity of their research experiences.

Students also reported that experiencing Iteration (36.5%) contributed to their perception that they were participating in real research. Many students explained that experiencing Iteration throughout the multiweek lab experience allowed them to understand that scientific research was not necessarily a quick and easy process:

> In the [regular] lab typically we would just spend a couple of hours studying something, but real research is done over time. I realize now it can be very repetitive [in response to question 6].

Experiencing Scientific Practices, using scientific tools or participating in the scientific process, was a frequently cited (36.5%) explanation of why students felt their experiences were authentic. Statements that this code applied to were often straightforward and frequently alluded to the scientific method or listed scientific activities, as in the following quote:

> I think we conducted real scientific research in this class because we ran a real experiment like researchers do. We follow-up step-by-step on the rules needed for an experiment like: creating a hypothesis, setting up a control, following up on the parameter every week and analyzing data [in response to question 4].

Students also discussed that their perceptions of research authenticity were bolstered by experiencing what we labeled Autonomy (22%), which (in addition to autonomy) could include a sense of project ownership or creative license. In their discussions of autonomy, students often described an increased appreciation for scientific research and for the CURE itself, as they felt they were expected to think more independently and realized that there was not always one "right" answer both specifically in their course and in science in general. For example:

> The main perception that changed was the amount of 'freedom' and 'creativity' you're allowed to have when doing scientific research. I thought that you would have stricter guidelines to conducting experiments. However, as a researcher the way you conduct your experiments is entirely up to you, and there are many different ways to determine the answer you are looking for. I was happy to discover that scientific research encourages creativity [in response to question 6].

TABLE 5. Coded elements that contributed to students' perceptions that their research experience was "real" (authentic)

"Real" research codes	Example quote[a]
Failure: Experiencing failure or setbacks	I always thought scientific research always runs smoothly or everything usually goes as planned. This made me realize that it's a lot of work to conduct scientific research and [experiments] don't run perfectly. There are always going to be some flaws or some negative outcomes. (*In response to Question 6*)
Iteration: Repeating experiments, or doing the experiment over a period of weeks	I prefer [the CURE] lab because it is more like real research... In this format we are able to trace the experiment for weeks and we have this opportunity to figure out the problems, and finally the [end] result is more reliable. (*In response to Question 8*)
Scientific practices: Using the practices, methods, tools, or processes of science	I have learned that scientific research is different than what I was expecting. I thought it was all theories and proving them. However, it's a technique and a deep research on identifying relevant data, and gathering it, and testing the hypothesis using a scientific method, and studying each change on the subject. (*In response to Question 6*)
Relevant Discovery: The potential for novel scientific discovery and/or the relevance of the project to the scientific community	To actually meet the person we're doing this research for really changes our perspectives. Being able to ask him questions on a personal level validates the point and purpose of why we're even doing it. (*In response to Question 1*)
Autonomy: Having autonomy, project ownership, or creative license (including in experimental design and interpreting results)	There are no real set guidelines [in research] since you are trying to "discover" something. You actually face trial and errors and try to find a solution to rectify this problem which was cool to see. It's great to actually use my own brain for once and try to figure out the data I am collecting and what it means. (*In response to Question 6*)
Collaboration: Working with classmates on their research project	I somewhat feel like I did [conduct real research] because I am working together with my teammates to figure out how to do a specific task in order to get the result we want to see. We all worked together to brainstorm and when our experiment failed, we would try to figure something else that could work better. (*In response to Question 4*)
"Successful" science: Producing data or results, experiencing success in experiments, or answering research questions	I do feel as if I have conducted real scientific research in this term of biology lab. The goal was to try to simulate an environment where the embryos would enter into diapause I, and my group was successful in doing so. Although having another species with the embryos might not be the exact and only reason that the embryos went into diapause I, it is a step closer to the right answer, or it may be part of the factors to the right answer. (*In response to Question 4*)

[a]Quotes have been lightly edited for grammar and concision. Question list is available in Table 2.

Relevant Discovery (29.7%) and Collaboration with teammates (12%) also contributed to students' perceptions that research felt real. While students frequently mentioned the faculty researcher whose research program was the focus of the CURE, these were almost exclusively about the increased awareness of the potential for Relevant Discovery within the CURE:

> I appreciated when [the faculty researcher] went into greater detail about the relevance of the experiment. It's easy to just focus on the basic aspects of the experiment like they're just a one-shot lab intended to teach a concept. Placing this in a larger picture with a large, unanswered question was cool [in response to question 1].

We therefore coded these instances as "Relevant Discovery" rather than "Collaboration." Interacting with the faculty researcher seemingly had a powerful effect on student discussions of Relevant Discovery—64% of students who indicated that Relevant Discovery made the course feel like real research connected this at least in part to interacting with the faculty researcher. While Collaboration as defined in the CURE literature can include collaboration with teammates, researchers, and instructors (Auchincloss *et al.*, 2014), students did not reflect on collaborating with their lab instructors, and we therefore coded Collaboration exclusively when students indicated working with their teammates:

> I have come to the realization that research is often a team effort, and collaboration is one of the most important parts [in response to question 6].

Finally, only 3% of students described that experiencing "Successful" Science was the reason that their lab experience felt like "real research," as exemplified in the following quote:

> I feel like we did [conduct real scientific research]; we actually got several embryos to enter diapause so that was a win! Not everything was ruined by the embryonic deaths [in response to question 4].

Very few students ended the term with sufficient sample sizes to conduct statistical analyses that could robustly address their hypotheses, so it is unsurprising that few students discussed the success of their experiments in lab.

Similar Experiences Can Have Variable Impacts on Student Perceptions of Research Authenticity

We coded the same set of reflection responses with an eye for identifying aspects of the experience that may have detracted from the perceived authenticity of the CURE. These statements were much less prevalent, and on average, students described only 0.4 unique course elements (SD = 0.6) that made their experiences feel inauthentic. Student critiques of how their experiences deviated from an authentic research experience

were thoughtful and often fair assessments of the limitations of the CURE; for example, 9% of students discussed the lack of time to continue their experiments:

> It is unfortunate that we do not have a longer period of time for data collection. I feel that more time would allow for more conclusive results to be drawn, due to the number of experimental conditions that had to be changed and the low rate of survival experienced with embryos [in response to question 4].

Other elements students described as making the course feel less authentic included a lack of significant results (12%) and a relative lack of autonomy (7%; Table 6). Many of the reasons provided for why their experiences felt inauthentic mirrored reasons other students cited as authentic research elements (Figure 2). For example, while most students (59%) interpreted their experiences of failure as a natural part of research, 4% of students interpreted those same experimental failures as indicators that they had *not* participated in "real" research:

> Overall, I feel like I did not conduct real scientific research this term … For Experiment 1, 6 embryos were alive, and potentially in diapause. However, in week 2, they all died. With Experiment 2, after adjusting our treatment, all 28 embryos died. With this, our group could not perform any type of statistical test [in response to question 4].

Similarly, while many students perceived that their opportunities for Iteration (36%), use of Scientific Practices (35%), and Autonomy (22%) over their experiments made their experiences feel real, other students felt that their experience was not real because of insufficient Iteration (9%), use of Scientific Practices (4%), or Autonomy (7%).

DISCUSSION

CURE Students Perceive Higher Levels of Collaboration, Relevant Discovery, and Iteration

In this study, we first aimed to quantitatively compare student perceptions of specific authentic research elements in two different lab types: a CURE and an inquiry-based course. We measured student perceptions of Collaboration, Relevant Discovery, and Iteration. Though both CURE and inquiry students recognized high levels of these elements in their laboratory courses, CURE students perceived statistically higher levels of each element. Notably, the effect size for the difference between perceived Collaboration was relatively small, which makes sense, given that CURE and inquiry students both collaborated in small and similarly structured groups. If we consider that Collaboration, Relevant Discovery, and Iteration are components of an authentic research experience, these results offer some clarification to the few previous attempts to compare CURE and inquiry student perceptions of research authenticity in the literature. Rowland and colleagues (2016) found that both CURE and inquiry students described their experiences as "real," and our results suggest that, while this may be true, CURE students may still perceive higher degrees of authenticity in their laboratory experiences. This supports recent findings that CURE students agree more strongly with the statement that they conducted scientific research in their lab courses than students who experienced lab curricula that lacked Relevant Discovery (Cooper *et al.*, 2019).

CURE students in particular reported higher perceived levels of Iteration compared with the inquiry students, which is notable, given that both CURE and inquiry students have two experimental iterations. CURE students conducted their experiments over a longer period of time (6 weeks compared with 4 weeks), and the instructors and faculty killifish researcher (J.E.P.) worked with CURE students to plan their second experimental iteration with great intentionality to help students build

TABLE 6. Coded elements that contributed to students' perceptions that their research experience was not "real" (inauthentic)

Research was not "real" codes	Example quote[a]
Insignificant Results: Lack of importance of results (possibly due to small sample size/lack of replicates)	It [the CURE] definitely had more scientific authenticity to it than any other bio lab, but I still felt like we didn't have enough data to be relevant.
Lacked time for iteration: Insufficient time to repeat experiments, to confirm results, etc.	The only part I feel that we are missing from this overall scientific research is time since we aren't going to be able to see what happens with the data we collected.
Lacked autonomy: Lack of student autonomy or control over the experiment	It [the CURE] is in a very controlled setting so in a way it does not feel real.
Experiments/tools too simple: Lack of sophisticated techniques or instruments	I feel like the methods we used were not very advanced and didn't give us enough precision to determine any real reasoning behind why embryos go into diapause.
Failure: Experiments failing, or not producing data that could conclusively address the research question	For the most part [our research felt real], but also not really, since our comparison group's eggs all died, so we don't really have anything super conclusive yet.
Outcomes were already known: Perception that instructors knew what was "supposed" to happen	I feel like the professors already know the outcomes to our experiments because they have probably done them before.

[a]Quotes have been lightly edited for grammar and concision. All quotes were reflections in response to Question #4: *Do you feel that you conducted real scientific research this term?"*
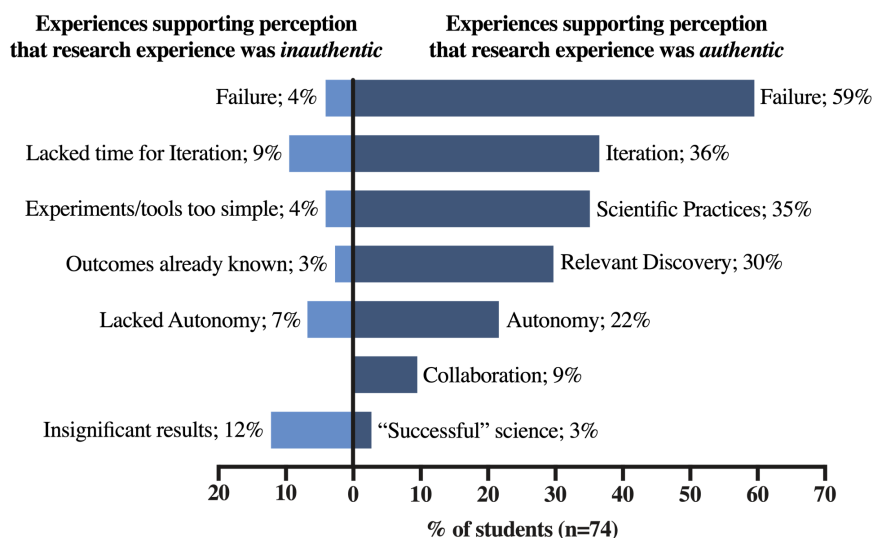
FIGURE 2. Parallel factors contributing to the CURE student's perceptions that their research experience was authentic or inauthentic.

upon what they had learned from their first experimental attempt. Although CURE students scored higher than inquiry students on each item within the LCAS Iteration subscale, CURE students reported particularly high perceived opportunities to revise their analyses and presentations based on feedback (LCAS items I5 and I6; Supplemental Material, Appendix 4). CURE students did not have more opportunities for formal formative feedback, so these items may reflect potentially increased attention that CURE instructors gave to their students in iterating their experiments and interpreting their results. Due to these efforts, CURE students may have had a better understanding and placed more value on the opportunity for Iteration. This aligns with previous evidence that students in research-based courses may develop an improved understanding of the nature of science: a large-scale qualitative study found that undergraduates in traditional, inquiry, and research-based labs had similar basic conceptions of different aspects of the nature of science, but inquiry and research-based students were able to articulate their understanding of the nature of science with respectively increased sophistication (Russell and Weaver, 2011).

Experiencing Elements of the Process of Science within the Context of "Participatory" Research May Be Key to Student Perceptions of Research Authenticity

In Table 1, we propose that CUREs align with a "participatory" model of authentic research in which Relevant Discovery and pursuing the "products" of science are prioritized (Auchincloss *et al.*, 2014; Barab and Hay, 2001; Spell *et al.*, 2014). However, in analyzing CURE student reflections to understand how different aspects of their experiences impact their perceptions of research authenticity, we found that Relevant Discovery was only the fourth most prevalent factor that students reported contributing to the authenticity of their experiences. Rather, students most commonly described experiencing Failure as making their experiences feel "real." We define failure as the inability to achieve a specific goal: these experiences were more serious than easily rectified errors but also did not discourage

students from persisting in redesigning and repeating their CURE research projects (Henry *et al.*, 2019). The top three elements contributing to perceived authenticity (experiencing Failure, Iteration, and Scientific Practices) all are arguably "process" of science elements that could occur in either simulated (inquiry) or participatory (CURE) models of authentic research. However, student reflections often indicated that experiences of failure were powerful in the CURE because the lack of a predetermined experimental scheme and expectations to confirm a previously tested hypothesis made failure feel inherently acceptable in the course. While teaching the CURE, instructors deliberately held discussions with students about how their experiences of challenges and failure are experiences inherent to scientific research, so it is unsurprising to see this perception mirrored in the student reflections. Additionally, the lack of performance-based goals and normalization of failure within our CURE likely served to reduce student stress and encourage "adaptive academic coping" behaviors, which are predicted to foster resiliency and challenge-seeking behaviors in students (Henry *et al.*, 2019). CURE student reflections also displayed an understanding that collecting reliable data was an important contribution to addressing the novel killifish research question and that time, patience, and iteration are necessary components to producing reliable data.

These findings mirror those of Gin *et al.* (2018), who found that students in a "high-challenge" course in which CURE students mostly failed to "successfully" answer their research questions responded more positively to their repeated experiences of iteration than students in a parallel "low-challenge" course. Further, students in the high-challenge course reported experiencing the same outcomes as students who did not experience as much failure or iteration in the course, indicating that failure and iteration did not detract from the positive benefits of CURE participation. Rather, they found that the context of Relevant Discovery that was inherent to the course design motivated students who experienced challenges and likely elevated the perceived importance of Iteration for students (Gin *et al.*, 2018).

From these observations, we propose that, while Failure and Iteration could occur in either simulated (inquiry) or participatory (CURE) models of authentic research, these elements are particularly powerful for students who are engaged in a participatory model of research and experience Relevant Discovery. This hypothesis is supported by the survey data: CURE students reported higher levels of Iteration compared with inquiry students, despite both curricula offering opportunities for Iteration. In other words, the context of the CURE may promote student buy-in to the authenticity of their research experiences to a greater extent than "simulating" research in an inquiry course, though CURE students may still prioritize the "process" of science elements that are common to both CURE and inquiry courses when considering the authenticity of their research

experiences. By increasing student buy-in during research-based courses through experiences of Iteration, Failure, and Relevant Discovery, we may also increase student engagement in learning and performance (Cavanagh *et al.*, 2016).

Alignment of Student and Expert Perceptions of Authentic Research

We compared how student perceptions of which research elements made their experiences feel "real" with both the CURE constructs and expert definitions of real research (Table 7). Although failure was the top explanation students gave for why their research felt real, this research element is not present in either the expert definitions of research (compiled in Rowland *et al.*, 2016) or in the originally proposed CURE constructs (Auchincloss *et al.*, 2014). Failure may therefore be a critical and previously underestimated experience for undergraduates in research-based courses. In light of this, researchers and curricular designers may want to focus their attention on framing and studying experiences of failure, as colleagues have begun to with the Failure as a Part of Learning: A Mindset Education Network (FLAMEnet) initiative (Heemstra *et al.*, n.d.).

The remaining elements that students identified as components of an authentic research experience were also recognized in at least one source by experts as an authentic research element. Iteration is included within the original CURE framework, but not in the expert definitions compiled by Rowland *et al.* (2016). Use of multiple Scientific Practices, Relevant Discovery, and Collaboration were elements of authenticity agreed on by students and experts—these elements were present in Rowland's compiled expert definitions of research authenticity and in the original CURE constructs. Finally, nearly a quarter of the CURE students discussed the importance of student autonomy, ownership, or creative license in supporting the perceived authenticity of their experiences. Although ownership is not a part of the original CURE framework, there have been several previous suggestions that ownership or autonomy is important in creating an authentic research experience for students (Barab and Hay, 2001; Rahm *et al.*, 2003; Hanauer *et al.*, 2012; Rowland *et al.*, 2016; Wald and Harland, 2017), particularly in CUREs (Hanauer and Dolan, 2014; Gin *et al.*, 2018; Cooper *et al.*, 2019).

Intriguingly, many of the experiences that the majority of students reported contributing to their perceptions of authentic research triggered the opposite conclusion for a minority of students. For example, while most found that failure and the opportunity for iteration made the experience feel more real, a few reasoned that their failures and the lack of time for increased iteration were what detracted from the authenticity of their research experiences. Recent research has suggested that, while failure can be a productive experience for undergraduates in a CURE (Gin *et al.*, 2018), and CURE instructors view opportunities for students to deal with failure as beneficial for students (Shortlidge and Brownell, 2016), experiencing failure in research can also be a factor in exacerbating depression for apprentice-based undergraduate researchers (Cooper *et al.*, 2020). We join Cooper *et al.* (2020) in hypothesizing that student researchers' variable perceptions of failure, and of other elements in our CURE, could be due to student mindset: students with a growth mindset may interpret challenges as productive learning experiences, while students with a fixed mindset tend to give up easily and respond negatively to setbacks (Dweck, 2008; Henry *et al.*, 2019). In our CURE, instructor-led discussions about the normalcy of failure in scientific research likely contributed to the majority of students recognizing failure as an experience to be expected when conducting scientific research. Because of the variable ways that students may interpret these experiences, instructors should be deliberate in normalizing failure and carefully framing these experiences for their students to promote productive student learning experiences and a growth mindset.

Student Reflections Provide Content Validity Evidence Supporting the CURE Framework

The CURE framework as proposed by Auchincloss and colleagues (2014) was derived through discussions with a small group of people experienced in CURE instruction and assessment who aimed to outline the elements necessary to engage students in research within the space of a course. To our knowledge, the degree to which the CURE framework elements lead to a perceived "authentic" experience for undergraduate students has not been externally validated by the

TABLE 7. Comparison of student and expert perspectives of authentic research design elements in the classroom[a]

Factor	Percent of students reporting factor contributed to "real research" perception (*n* = 74)	Equivalent themes in expert definitions of research authenticity elements (Rowland *et al.*, 2016)	CURE constructs (Auchincloss *et al.*, 2014)
Failure	59	×	×
Iteration	36	×	√
Use of multiple scientific practices	35	√[b]	√
Relevant Discovery	30	√[c]	√[d]
Autonomy	22	√[e]	×
Collaboration	9	√[f]	√

[a]An "×" indicates that the factor was not present in the expert definitions or CURE constructs, while a √ indicates that the factor was present in the expert definitions or CURE constructs.
[b]Described in Rowland *et al.*, 2016, as "Experience of what scientists "do" (practices), how science is done, and what science "is""; Experimental design; Communication; Data analysis".
[c]Described in Rowland *et al.*, 2016, as "Results are novel/publishable/contribute to existing research; Audience (real problem); Outcome is unknown to all".
[d]Described in Auchincloss *et al.*, 2014, as separate constructs ("Discovery" and "Broader Relevance").
[e]Described in Rowland *et al.*, 2016 , as "Ownership/personal relevance to students; Critical thinking".
[f]Described in Rowland *et al.*, 2016, as "Peer teamwork".

target population. Through our work, we are able to test whether the aspects that make a CURE feel like "research" to the target population (undergraduate students) converge with the expert-defined CURE framework. Though our reflection questions did not directly probe students about the CURE framework elements, we found that each element—Iteration, use of multiple Scientific Practices, Collaboration, and Relevant Discovery—was present in student descriptions of what made their research experience feel authentic. These data indicate that intentionally scaffolding each of these elements, in conjunction with providing students with opportunities for Failure and Autonomy, will best support CURE students in perceiving that they are participating in real research.

Limitations

There are several limitations to this study, in particular with regard to our attempts to compare the experiences of inquiry and CURE students. The inquiry and CURE courses occurred concurrently and engaged students from the same student population, but there were some differences between the curricula that could have variable impacts on the perceptions that CURE and inquiry students had of their experiences. The CURE and inquiry project study organisms were very different—CURE students worked with fish and their embryos, while inquiry students worked with sowbugs. Though we do not have data on this, we anecdotally have observed a range of student reactions to working with both of these study organisms, including disgust and boredom (especially sowbugs), squeamishness and excitement to be working with living organisms (both sowbugs and killifish), and enthrallment (particularly for killifish). These perceptions and attitudes may affect student interest and motivation in engaging with the course (Hidi and Renninger, 2006), which could ultimately be reflected in the way students responded in the survey and reflection questions they completed for this study.

CURE students spent an additional 2 weeks on their work, and the additional time likely allowed the GTAs to spend more time providing in-class formative feedback to their students. In combination with the study reflection questions, this could have aided the CURE students in thinking more deeply about their experiences. Our qualitative data were self-reported by our student participants through reflection questions that would be read by their GTAs, and this context could potentially lead to bias in student responses, though we tried to mitigate this by making it clear that the reflection questions were not graded for content. While the sample size from our CURE students is sufficient to provide us with extensive qualitative information, we had limited resources to scale up the CURE to more laboratory sections, and therefore lack the sample size needed to conduct more statistically appropriate quantitative comparisons between CURE and inquiry students. Further, while we initially chose to focus our qualitative data collection on CURE students who would be able to report their experiences with both "process" and "product" of science elements, in retrospect, we would have extended this study by administering similar reflection questions to both CURE and inquiry students to further explore the differences and similarities in how CURE and inquiry students operationalize research authenticity in their classrooms. Our plans to expand data collection in subsequent terms to increase our statistical power and comparisons between CURE and

inquiry students were thwarted by 1) a collapse of our killifish study system in Spring 2019 and 2) the COVID-19 pandemic in Spring 2020.

Finally, our data are only representative of one introductory biology university population and may not be representative of student experiences in other institutional contexts, particularly given the relatively high proportions of transfer, non–traditional aged, and postbaccalaureate students within our study population. Though a previous study with our student population found that student age or postbaccalaureate status did not have much impact on student perceptions of the classroom (Shortlidge *et al.*, 2019), older students are more likely to endorse learning-oriented rather than performance goals and are therefore likely to have a stronger growth mindset and resiliency to failure (Dweck, 2008; Eppler *et al.*, 2000). The relatively high proportion of non–traditional aged students within our student population is therefore an unexplored potential explanation for our students' positive reactions to failure in the CURE.

CONCLUSION

Overall, we found that most students who participated in a novel killifish CURE believed they were indeed participating in real research, and we found significant overlap between expert and student explanations of what constitutes an authentic research experience. Interestingly, CURE students largely attributed experiences of failure and iteration to why they felt they had participated in real research. Therefore, if instructors of discovery-based courses aim for students to believe that they are participating in real research, they may want to consider how to leverage and positively facilitate these experiences in curricular design to promote student buy-in.

As educators and researchers, we often believe that research experiences are beneficial for students. However, we do not know how important it is for students to believe they are experiencing real research in order to reap the benefits of research participation. We propose that future research explore whether students need to buy into the authenticity of their research experiences to benefit from their exposure to research. Further, if students do need to believe that their research experiences are authentic in order to experience the benefits of research participation, do their perceptions of research authenticity need to align with the expert expectations and beliefs of what makes a classroom research experience authentic? This work contributes to our growing understanding of student perceptions of evidence-based teaching and of the value of how discovery-based curricula can offer more equitable access to authentic research experiences.

findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

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

Bakshi, A., Patrick, L. E., & Wischusen, E. W. (2016). A framework for implementing course-based undergraduate research experiences (CUREs) in freshman biology labs. *American Biology Teacher*, *78*(6), 448–455. https://doi.org/10.1525/abt.2016.78.6.448

Ballen, C. J., Blum, J. E., Brownell, S., Hebert, S., Hewlett, J., Klein, J. R., … & Cotner, S. (2017). A call to develop course-based undergraduate research experiences (CUREs) for nonmajors courses. *CBE—Life Sciences Education*, *16*(2), mr2. https://doi.org/10.1187/cbe.16-12-0352

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.

Barab, S. A., & Hay, K. E. (2001). Doing science at the elbows of experts: Issues related to the science apprenticeship camp. *Journal of Research in Science Teaching*, *38*(1), 70–102. https://doi.org/10.1002/1098-2736(200101)38:1<70::AID-TEA5>3.0.CO;2-L

Barbera, J., & VandenPlas, J. R. (2011). All assessment materials are not created equal: The myths about instrument development, validity, and reliability. In *Investigating classroom myths through research on teaching and learning* (Vol. 1074, pp. 177–193). Washington DC: ACS Symposium Serie, American Chemical Society. https://doi.org/10.1021/bk-2011-1074.ch011

Beck, C. W., & Blumer, L. S. (2016). Alternative realities: Faculty and student perceptions of instructional practices in laboratory courses. *CBE—Life Sciences Education*, *15*(4), ar52. https://doi.org/10.1187/cbe.16-03-0139

American Association for the Advancement of Science. (2011). *Vision and change in undergraduate biology education: A call to change*. Washington, DC.

Brownell, S. E., & Kloser, M. J. (2015). Toward a conceptual framework for measuring the effectiveness of course-based undergraduate research experiences in undergraduate biology. *Studies in Higher Education*, *40*(3), 525–544. https://doi.org/10.1080/03075079.2015.1004234

Brownell, S. E., Kloser, M. J., Fukami, T., & Shavelson, R. J. (2013). Context matters: Volunteer bias, small sample size, and the value of comparison groups in the assessment of research-based undergraduate introductory biology lab courses. *Journal of Microbiology & Biology Education*, *14*(2), 176–182. https://doi.org/10.1128/jmbe.v14i2.609

Buck, L. B., Bretz, S. L., & Towns, M. H. (2008). Characterizing the level of inquiry in the undergraduate laboratory. *Journal of College Science Teaching*, *38*(1), 52–58.

Cavanagh, A. J., Aragón, O. R., Chen, X., Couch, B. A., Durham, M. F., Bobrownicki, A., … & Graham, M. J. (2016). Student buy-in to active learning in a college science course. *CBE—Life Sciences Education*, *15*(4), ar76. https://doi.org/10.1187/cbe.16-07-0212

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., Gin, L. E., Barnes, M. E., & Brownell, S. E. (2020). An exploratory study of students with depression in undergraduate research experiences. *CBE—Life Sciences Education*, *19*(2), ar19. https://doi.org/10.1187/cbe.19-11-0217

Cooper, K. M., Soneral, P. A. G., & Brownell, S. E. (2017). Define your goals before you design a CURE: A call to use backward design in planning course-based undergraduate research experiences. *Journal of Microbiology & Biology Education*, *18*(2). https://doi.org/10.1128/jmbe.v18i2.1287

Corwin, L. A., Graham, M. J., & Dolan, E. L. (2015a). 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. (2015b). The Laboratory Course Assessment Survey: A tool to measure three dimensions of research-course design. *CBE—Life Sciences Education*, *14*(4), ar37.

Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Thousand Oaks, CA: Sage.

Dalgaard, P. (2008). *Introductory statistics with R*. New York, NY: Springer.

Domin, D. S. (1999). A review of laboratory instruction styles. *Journal of Chemical Education*, *76*(4), 543. https://doi.org/10.1021/ed076p543

Dweck, C. S. (2008). *Mindset: the new psychology of success*. New York, NY: Ballantine Books.

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

Eppler, M. A., Carsen-Plentl, C., & Harju, B. L. (2000). Achievement goals, failure attributions, and academic performance in nontraditional and traditional college students. *Journal of Social Behavior & Personality*, *15*(3), 353–372.

Esparza, D., Wagler, A. E., & Olimpo, J. T. (2020). Characterization of instructor and student behaviors in CURE and Non-CURE learning environments: Impacts on student motivation, science identity development, and perceptions of the laboratory experience. *CBE—Life Sciences Education*, *19*(1), ar10. https://doi.org/10.1187/cbe.19-04-0082

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences USA*, *111*(23), 8410–8415. https://doi.org/10.1073/pnas.1319030111

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

Govindan, B., Pickett, S., & Riggs, B. (2020). Fear of the CURE: A beginner's guide to overcoming barriers in creating a course-based undergraduate research experience. *Journal of Microbiology & Biology Education*, *21*(2). https://doi.org/10.1128/jmbe.v21i2.2109

Graham, M. J., Frederick, J., Byars-Winston, A., Hunter, A.-B., & Handelsman, J. (2013). Increasing persistence of college students in STEM. *Science*, *341*(6153), 1455–1456. https://doi.org/10.1126/science.1240487

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

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

Hancock, G. R., Stapleton, L. M., Mueller, R. O., Stapleton, L. M., & Mueller, R. O. (2018). *The reviewer's guide to quantitative methods in the social sciences*. New York, NY: Routledge. https://doi.org/10.4324/9781315755649

Heemstra, J. M., Corwin, L. A., Charkoudian, L., Le, B., Henry, M. A., & Shorter, S. (n.d.). *Failure as a part of learning: A Mindset Education network*. Retrieved July 13, 2020, from https://qubeshub.org/community/groups/flamenet

Henry, M. A., Shorter, S., Charkoudian, L., Heemstra, J. M., & Corwin, L. A. (2019). FAIL is not a four-letter word: A theoretical framework for exploring undergraduate students' approaches to academic challenge and responses to failure in STEM learning environments. *CBE—Life Sciences Education*, *18*(1), ar11. https://doi.org/10.1187/cbe.18-06-0108

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational psychologist*, *41*(2), 111–127. https://doi.org/10.1207/s15326985ep4102_4

Hofstein, A., & Lunetta, V. N. (2004). The laboratory in science education: Foundations for the twenty-first century. *Science Education*, *88*(1), 28–54. https://doi.org/10.1002/sce.10106

Komperda, R. N., Hosbein, K., & Barbera, J. (2018a). Evaluation of the influence of wording changes and course type on motivation instrument functioning in chemistry. *Chemistry Education Research and Practice*, *19*(1), 184–198. https://doi.org/10.1039/C7RP00181A

Komperda, R., Pentecost, T. C., & Barbera, J. (2018b). Moving beyond alpha: A primer on alternative sources of single-administration reliability evidence for quantitative chemistry education research. *Journal of Chemical Education*, *95*(9), 1477–1491. https://doi.org/10.1021/acs.jchemed.8b00220

Laursen, S., Hunter, A.-B., Seymour, E., Thiry, H., & Melton, G. (2010). *Undergraduate research in the sciences: Engaging students in real science*. San Fransisco, CA: Wiley.

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

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

National Research Council. (1996). *National Science Education Standards*, Washington, DC: The National Academies Press. https://doi.org/10.17226/4962

O'Connor, C., & Joffe, H. (2020). Intercoder reliability in qualitative research: debates and practical guidelines. *International Journal of Qualitative Methods*, *19*. https://doi.org/10.1177/1609406919899220

Peters, G.-J. Y. (2018). *userfriendlyscience: Quantitative analysis made accessible*. https://doi.org/10.17605/osf.io/txequ

President's Council of Advisors on Science and Technology. (2012). *Engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics*. Washington, DC: U.S. Government Office of Science and Technology.

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

Rocabado, G. A., Komperda, R., Lewis, J. E., & Barbera, J. (2020). Addressing diversity and inclusion through group comparisons: A primer on measurement invariance testing. *Chemistry Education Research and Practice*, *21*, 969–988. https://doi.org/10.1039/D0RP00025F

Rosenthal, R. (1965). The volunteer subject. *Human Relations*, *18*(4), 389–406. https://doi.org/10.1177/001872676501800407

Rosseel, 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

RStudio Team. (2019). *RStudio: Integrated development for R*. Boston, MA: RStudio, Inc. www.rstudio.com/

Russell, C. B., & Weaver, G. C. (2011). A comparative study of traditional, inquiry-based, and research-based laboratory curricula: Impacts on understanding of the nature of science. *Chemistry Education Research and Practice*, *12*(1), 57–67. https://doi.org/10.1039/C1RP90008K

Saldana, J. (2015). *The coding manual for qualitative researchers*. Thousand Oaks, CA: Sage.

Shortlidge, E. E., Bangera, G., & Brownell, S. E. (2016). Faculty perspectives on developing and teaching course-based undergraduate research experiences. *BioScience*, *66*(1), 54–62. https://doi.org/10.1093/biosci/biv167

Shortlidge, E. E., Bangera, G., & Brownell, S. E. (2017). Each to their own CURE: Faculty who teach course-based undergraduate research experiences report why you too should teach a CURE. *Journal of Microbiology & Biology Education*, *18*(2). https://doi.org/10.1128/jmbe.v18i2.1260

Shortlidge, E. E., & Brownell, S. E. (2016). How to Assess Your CURE: A Practical Guide for Instructors of Course-Based Undergraduate Research Experiences. *Journal of microbiology & biology education*, *17*(3), 399–408. https://doi.org/10.1128/jmbe.v17i3.1103

Shortlidge, E. E., Rain-Griffith, L., Shelby, C., Shusterman, G. P., & Barbera, J. (2019). Despite similar perceptions and attitudes, postbaccalaureate students outperform in introductory biology and chemistry courses. *CBE—Life Sciences Education*, *18*(1), ar3. https://doi.org/10.1187/cbe.17-12-0289

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

Sundberg, M. D., Armstrong, J. E., & Wischusen, E. W. (2005). A reappraisal of the status of introductory biology laboratory education in US colleges & universities. *American Biology Teacher*, *67*(9), 525–529.

Thompson, S. K., Neill, C. J., Wiederhoeft, E., & Cotner, S. (2016). A model for a course-based undergraduate research experience (CURE) in a field setting. *Journal of Microbiology & Biology Education*, *17*(3), 469–471. https://doi.org/10.1128/jmbe.v17i3.1142

Vereijken, M. W. C., van der Rijst, R. M., de Beaufort, A. J., van Driel, J. H., & Dekker, F. W. (2016). Fostering first-year student learning through research integration into teaching: Student perceptions, beliefs about the value of research and student achievement. *Innovations in Education and Teaching International*, *55*(4), 425–432. https://doi.org/10.1080/14703297.2016.1260490

Vereijken, M. W. C., van der Rijst, R. M., Van Driel, J. H., & Dekker, F. W. (2019). Authentic research practices throughout the curriculum in undergraduate medical education: Student beliefs and perceptions. *Innovations in Education and Teaching International*, *57*(5), 532–542. https://doi.org/10.1080/14703297.2019.1674680

Wald, N., & Harland, T. (2017). A framework for authenticity in designing a research-based curriculum. *Teaching in Higher Education*, *22*(7), 751–765. https://doi.org/10.1080/13562517.2017.1289509