Impact of Authentic Course-Based Undergraduate Research Experiences (CUREs) On Student Understanding in Introductory Biology Laboratory Courses

Joseph LaForge and Erika C. Martin

ABSTRACT
We present data displaying course-based undergraduate research experiences (CUREs) effectiveness in providing authentic cutting-edge research experiences to undergraduates, which both private and government organizations recognize as essential. A total of 68 students were enrolled in this research with 30 students being in a traditional laboratory course and 18 students participating in the CURE implemented laboratory. Results from mid- and postsemester surveys were compared to assess knowledge and attitude. Knowledge showed no change; however, students who experienced the CURE responded with increased enjoyment, strong feelings of scientific contribution, and high project ownership, and overall they were more confident in research than their non-CURE peers.

Key Words: student attitude; scientific identity; attitude toward science; research in classroom.

Introduction
A discussion of undergraduate education in 1953 led experts at the National Science Foundation (NSF) to fund and provide sites for universities to increase undergraduate research (Kinkead, 2012). NSF interest was followed by Professor Margaret MacVicar at the Massachusetts Institute of Technology, where in 1969 MacVicar founded the Undergraduate Research Opportunities Program (UROP). This program rapidly increased in size and diversity (Cohen & MacVicar, 1998). A separate entity, the Counsel of Undergraduate Research (CUR), was formed in 1978 by professors at liberal arts colleges and stimulated nationwide promotion of undergraduate research. In 1986, NSF began its Research Experiences for Undergraduates (REU) program (Kinkead, 2012). REU involves students in research promoted and designed by NSF. However, a study of 3400 STEM students showed that only 0.06% were sponsored by government organizations, such as the NSF, NASA, and the National Institute of Health (Russell et al., 2007). Precisely 99.94% of students were sponsored by institutions, states, or held no sponsorship. Thus, the vast majority of research opportunities are not funded by federal government organizations (Russell et al., 2007). An increasing demand for undergraduate research is not being met, and of the paucity of research opportunities, most are not coming from federal support; this requires a solution. One possible solution to increase student involvement is to incorporate research directly into required classes. Inclusion of entire classes in self-guided investigation is different from traditional laboratories because the research is semester-long and is of relevance to students’ own interests. Investigation creates opportunities for student involvement while learning remains the teacher’s highest priority. Science professors identify a need for revitalization of traditional lecture-based courses at the university level with inquiry and investigation, and this revitalization should be considered extremely relevant (Sundberg et al., 2005; Weaver et al., 2008). Student engagement in authentic research has significant positive impacts on development of science literacy and reasoning skills (Holt et al., 1969; Sundberg et al., 2005). Authentic research experiences also increase student retention in the sciences (Espinosa, 2011; Linn et al., 2015). Science students who enroll in laboratory sessions with traditional laboratory curricula do not show these same increases of retention (Wei & Woodin, 2011). Thus, research is an important component to an undergraduate science education. Many research experiences for students occur in the laboratory of individual faculty and are accessible to few students.

Course-Based Undergraduate Research Experiences (CUREs), only emerging in recent years (Auchincloss et al., 2014; Alkaher & Dolan, 2014; Dolan, 2016), apply investigative and inquiry-based teaching techniques to multitudes of students with diverse learning styles and backgrounds (Bangera & Brownell, 2014; Rowland et al., 2012). Course-based approaches such as CUREs offer research and inquiry to larger numbers of students (Weaver et al., 2008). Involvement in CUREs familiarizes students with the scientific process through exposure to creating and performing research. Studies of life science courses with CUREs show increases in content knowledge, self-efficacy, and persistence (Lopatto et al., 2008; Shaffer et al., 2014). Similar results of increased student retention and positive attitudes toward sciences are described in chemistry, geoscience, and physics (Ryan, 2014; Rodenbusch et al., 2016). CUREs provide students an environment to create their own discoveries.
through collecting and analyzing their own data. These discoveries can be relevant to the broader professional scientific community.

CUREs involve students in five activities: scientific practices, discovery, relevant work, collaboration, and iteration (Auchincloss et al., 2014). Students show increases in other areas as well, including evaluating literature, reasoning in a nonlinear, inquisitive fashion, and ability to replicate their methods (Dolan, 2016). Furthermore, CUREs create a stronger sense of project ownership and higher levels of persistence in the sciences than students who participate in traditional laboratory courses (Hanauer et al., 2012). Increased persistence has been shown in multiples studies (Drew & Triplett, 2008; Harrison et al., 2011; Hanauer et al., 2012; Bascom-Slack et al., 2012; Brownell et al., 2012; Jordan et al., 2014; Shaffer et al., 2014), and CUREs also help with career clarification (Drew & Triplett, 2008; Harrison et al., 2011; Shaffer et al., 2014).

We studied the impact of our novel CURE (described ahead) on student metrics of literacy, performance, and attitude using the aforementioned surveys. A particular concern for science educators is ensuring that students gain scientific literacy skills integral to informed decision making (National Academies of Sciences, Engineering, and Medicine, 2016). This project design immersed students in scientific practices. We sought to discover if student participation in the project would lead to scientific literacy and enjoyment of science. In addition, the intellectually ambitious nature of this project design required students to immerse themselves in scientific practices not generally obtained through traditional laboratory settings, including literature searches and evaluation of sources. Thus, we expected CURE students to have higher levels of scientific literacy (ability to identify valid scientific arguments, interpret graphical data to plausible descriptions, define strengths and weaknesses of an experiment, and deduce credibility of a scientific journal) and self-efficacy (confidence in one's own capacity to do well) than those in the traditional laboratory environment.

○ Methods

Inquiry-based learning is widely promoted to increase literacy (Gormally et al., 2009). One test that is widely used to quantify literacy and performance skills is the Test of Scientific Literacy Skills (TOSLS; see Gormally et al., 2012). TOSLS measures skills related to scientific literacy, such as evaluating the validity of sources, identifying valid scientific arguments, and identifying strengths and weaknesses in research design (Gormally et al., 2012). Revised through five semesters, TOSLS has been used in traditional lecture-based courses, as well as learner-centered courses, where the students were responsible for discovering new topics (Brickman et al., 2012; Gormally et al., 2012). TOSLS showed posttest scores that were significantly higher than pretest scores for project-based courses and traditional courses (Gormally et al., 2012). These results indicate that learning gains are similar in CURE and non-CURE classrooms. Research experiences also are found to have an impact on attitude (Chemers et al., 2011).

To quantify students’ attitudes toward science, we utilized Persistence in the Sciences (PITS; see Hanauer et al., 2016). PITS provides a reliable and valid way of measuring variables underpinning student science persistence (Hanauer et al., 2016). Development of PITS was informed by existing instruments, such as project ownership, self-efficacy, science identity, science community values, and networking (Hanauer et al., 2016). Validity checks that a given assessment truly measures what it claims to measure. Both construct and content validity are addressed. Construct validity ensures a relation to underlying theoretical concepts. Content validity ensures that test content is related to what is being tested. National reports, as well as faculty surveys, asserted content validity in PITS. Previous trials in four classes assessed the effectiveness of PITS with CUREs on student project ownership (attitude and content), science identity, and self-efficacy. Undergraduates’ attitude toward research should not be overlooked (Chang et al., 2011; Chemers et al., 2011; Harrison et al., 2011; Linn et al., 2015; Hanauer et al., 2016).

○ Demographics

CURE (n = 18) and non-CURE (n = 50) participants were students enrolled at Emporia State University (ESU) in an introductory biology laboratory—30% male and 70% female. ESU did not evaluate the spectrum of gender identities. Considering ethnicity, Whites composed 68%, those of two or more ethnicities 16%, Hispanics 9%, Blacks and/or African Americans 4%, and Asians 3%. No other categories were recorded. Freshmen accounted for 49%, sophomores 45%, and juniors 6%. Eleven different majors were represented: nursing (55%), biology (15%), and athletic training (8%). Other majors included art, business administration, chemistry, communication, health and human performance, political science, psychology, and undecided (combined enrollment = 22%).

ESU is a masters-granting institution in east central Kansas with approximately 3500 undergraduates enrolled and a total enrollment of approximately 3700 students. Three classes with sizes of 118, 21, and 29 were assessed during fall 2019. TOSLS and PITS were administered using hard copies during the regular lab period at midterm and in the last week of class. An online survey was given during the last week of class (Table 2).

○ Non-CURE & CURE Similarities & Differences

Both non-CURE and CURE students engaged in investigations through use of traditional laboratory equipment and traditional curriculum. Each week, a new biological topic was completed using a laboratory manual. Non-CURE students chose small-group (2–4 individuals) research projects. Groups frequently chose “safe” topics, such as assessing growth of bread mold or comparing plant growth. Conversely, the CURE laboratories involved students with engagement in investigation both during and outside of the scheduled laboratory time (described in detail ahead). Students asked the research question, created the hypothesis, collected data out-of-lab, reviewed literature, and functioned as a team.

○ CURE Description

Behavior researchers’ results suggest fishes (specifically the cleaner wrasse, Labroides dimidiatus) might have the capacity for self-awareness by passing the Mark Test (Kohda et al., 2019). Fish awareness is either self-recognition at its most basic level, or the more holistic “sense of self” that primates display (McCallum, 2019). Darters (Percidae) are small fishes native to North America. They have the physiological anatomy to see color. In some Ethostomus species,
vision is essential for mate choice. Both female splendid darter (E. barrenensis) and banded darter (E. zonale) rely on male coloration when choosing a mate (Williams & Mendelson, 2010). CURE students chose to study two species of darters: Ozark logperch (P. fulvitae) and slenderhead darter (P. phoxocephala), to see if they pass the mark test.

CURE students were given a nonprofessional self-awareness article (Buehler, 2019) and the peer-reviewed paper by Kohda et al. (2019) to read. Popular media is more accessible to students than peer-reviewed literature. Having students first read news coverage of a scientific article shows students that the topic is relevant, and it gives them social and societal context before reading peer-reviewed literature. CURE students preformed their observations in aquarium microcosms (see Martin et al., 2018) containing a single fish, including or excluding a mirror. Observations were performed both in-person and by recorded videos. CURE students used the data to see if fish displayed self-recognition. By semester’s end, students finalized their results and concluded the study by presenting a paper (Students’ CURE paper was recently published: Nelson et al., 2021; lesson design: LaForge and Martin, in review).

Our CURE explored whether native freshwater fishes were capable of self-recognition using the mark test. Research concepts, methods, plan of class investigation, literature review, data collection, writing the scientific article, and submitting the article for publication were all performed with direct input from CURE students.

### Results

#### Survey Results & Data Analysis

Data were analyzed and graphed in R (R Core Team, 2019), using libraries ggplot2 (Wickham, 2016), rcompanion (Mangiafico, 2016), car (Fox & Weisberg, 2019), lmeans (Lenth, 2016), and effsize (Torchiano, 2020).

PITS answers ranged from 1 to 5. A score of 1 indicated a student strongly disagreed with the statement and a score of 5 indicated a student strongly agreed. A score of 1 reflected negative experiences and 5 positive experiences across six categories (n= 36 total): Community (n = 4), Emotion (n = 6), Identity (n = 5), Networking (n = 5), Ownership (n = 10), and Self-Efficacy (n = 6). Median values, quartiles, and the percent of 1 and 5 responses were calculated and compared.

For TOSLS, which measures content and conceptual knowledge using multiple choice questions (N = 28). Possible scores are 0–28. Initial analysis indicated scores were abnormally distributed with a slight positive skew. A Tukey's ladder of powers assessment suggested a fourth-root transformation was necessary to meet normality (\( \lambda = 0.575, w = 0.971 \)). Levene's test for homogeneity found no heteroscedasticity between non-CURE and CURE classes (\( F_{1,129} = 0.356, P = 0.552 \)) nor individual courses (\( F_{2,128} = 0.291, P = 0.748 \)). We first conducted a two-way ANOVA comparing student scores (dependent variable) between the non-CURE and CURE courses (independent variable, two levels) across time points (independent variable, two levels, midterms and end of semester). We conducted a two-way ANOVA comparing individual courses (three levels) and time point; however, time was nonsignificant and removed from the model. Thus a one-way ANOVA was used to evaluate student scores across the three individual course sections. When significant differences were found, a Tukey's honestly significant difference (Tukey HSD) post hoc analysis was used to compare groups, and Cohen's \( d \) effect size was calculated (Cohen, 1988).

### Results

#### Survey Results & Data Analysis

Data were analyzed and graphed in R (R Core Team, 2019), using libraries ggplot2 (Wickham, 2016), rcompanion (Mangiafico, 2016), car (Fox & Weisberg, 2019), lmeans (Lenth, 2016), and effsize (Torchiano, 2020).

PITS answers ranged from 1 to 5. A score of 1 indicated a student strongly disagreed with the statement and a score of 5 indicated a student strongly agreed. A score of 1 reflected negative experiences and 5 positive experiences across six categories (n= 36 total): Community (n = 4), Emotion (n = 6), Identity (n = 5), Networking (n = 5), Ownership (n = 10), and Self-Efficacy (n = 6). Median values, quartiles, and the percent of 1 and 5 responses were calculated and compared.

For TOSLS, which measures content and conceptual knowledge using multiple choice questions (N = 28). Possible scores are 0–28. Initial analysis indicated scores were abnormally distributed with a slight positive skew. A Tukey's ladder of powers assessment suggested a fourth-root transformation was necessary to meet normality (\( \lambda = 0.575, w = 0.971 \)). Levene's test for homogeneity found no heteroscedasticity between non-CURE and CURE classes (\( F_{1,129} = 0.356, P = 0.552 \)) nor individual courses (\( F_{2,128} = 0.291, P = 0.748 \)). We first conducted a two-way ANOVA comparing student scores (dependent variable) between the non-CURE and CURE courses (independent variable, two levels) across time points (independent variable, two levels, midterms and end of semester). We conducted a two-way ANOVA comparing individual courses (three levels) and time point; however, time was nonsignificant and removed from the model. Thus a one-way ANOVA was used to evaluate student scores across the three individual course sections. When significant differences were found, a Tukey's honestly significant difference (Tukey HSD) post hoc analysis was used to compare groups, and Cohen's \( d \) effect size was calculated (Cohen, 1988).

### Results

Across all PITS categories, just 1% of CURE students answered with a negative attitude score of 1 compared to 14% of non-CURE students answering with 1. The percentage of students responding with 1, indicating a more negative experience, was consistently higher for non-CURE by at least 3%, but as great as 15%, depending on the category (Figure 1).

For CURE students, overall median response was a positive attitude score of 4. The percentage of students responding with the extremely negative attitude score of 1 was Community = 2%, Emotion = 0%, Identity = 6%, Networking = 8%, Ownership = 1%, and Self-Efficacy = 0%. For non-CURE students, overall median response was an indifferent attitude score of 3. Percentage of students responding with the extremely negative attitude score of 1 was Community = 8%, Emotion = 12%, Identity = 21%, Networking = 16%, Ownership = 13%, and Self-Efficacy = 3%.

Online survey questions included

- Did you enjoy your research project?
- Do you feel like the research project taught you about how real scientists conduct research?
- Do you believe the research will have a positive impact on your future in science?
- Do you believe your research was relevant to the scientific community?
- Would you want to be a part of a research team in the future?

Responses complemented results from PITS. The Yes answers indicate positive experiences. CURE students selected more positive responses for questions than non-CURE students (Table 1), in some instances 100% compared to 0%. CURE students responded Yes 73% of the time, No just 2% and Maybe/Unsure 25%. By contrast, non-CURE students responded Yes 30% of the time, No 42% and Maybe/Unsure 28% (Table 1).

For TOSLS, which measures content and conceptual knowledge, the initial two-way ANOVA comparing non-CURE and CURE student scores across time found no significant differences between course types (\( F_{1,127} = 0.060, P = 0.808 \)), between time points (\( F_{1,127} = 2.230, P = 0.138 \)), or in interactions between the two independent variables (\( F_{1,127} = 0.208, P = 0.649 \)). The one-way ANOVA comparing the final score of the three course sections found significant
differences among the three course sections ($F_{2,126} = 10.161$, $P < 0.001$, Figure 2). A Tukey HSD post hoc analysis indicated the two non-CURE course sections performed differently on the TOSLS ($\bar{x}_1 = 8.5$, range$_1 = 0 – 23$; $\bar{x}_2 = 13.0$, range$_2 = 5 – 28$) and the subsequent Cohen’s $d$ effect size analysis found this difference to be large ($d = 0.916$).

### Discussion

We compared student gains between CURE to the traditional laboratory. Assessments were given at the middle and end of the semester to measure gains in scientific knowledge, attitude, and identity. While only three headlines, much is discussed. Nested in the knowledge section is scientific literacy and student performance. Attitude has nested within it student enjoyment, contribution, research relevance, and project ownership. Last, we describe student identity, including confidence in performing science, community involvement, ability to perform scientific procedures, and student comfortability in a science laboratory.

### Knowledge

TOSLS did not show a change, either positive or negative, between CURE and non-CURE students. Previous literature varies on student performance. Some implementations suggest improved performance on general knowledge tests (Russell & Weaver, 2011; Ward, 2014). On the other hand, some implementations suggest similar learning gains between both groups of students (Wolkow, 2014). Our results suggest that students learn content equally from both traditional and CURE method. However, they learn other things, such as how to develop an investigation, differently. While results indicate equal content knowledge gained, CURE students were found to enjoy themselves more when involved in the discovery techniques offered.

### Attitude

Students’ enjoyment of research is increased in the CURE classroom. Greater contribution to the scientific community and greater research relevance is shown by students. Authentic research has long-lasting impacts on students’ attitude toward the sciences. Our results reinforce hypotheses described by Auchincloss et al. (2014), Alkaher & Dolan (2014), and Dolan (2016), which claim CUREs are effective in creating positive gains in attitude. Allowing for reflection of their progress, this CURE increased students’ project ownership. This project is a building block that creates memorable, positive impacts on students who perform research. Students may go on to be research project ambassadors who advocate for science through action.

### Identity

CURE students showed more confidence than non-CURE in identifying themselves as scientists. CURE students felt the research project taught them about how real scientists conduct research. Our data reinforces that with CURE, students are more comfortable with involvement in scientific communities. Identifying professionally as a scientist through conducting research is shown as beneficial by Seymour et al. (2004). Scientific identity includes its own type of performance. As described by Carlone & Johnson (2007), performance reassures students that they can be scientists by being able to perform scientific activities. Involvement in CURE shows that along with being able to perform scientific procedures, students desire to continue in the sciences (Hanauer et al., 2012), using their newly trained skills beyond the current course.

Research will continue by analyzing another CURE and non-CURE class for a relation between identity, performance, and attitude. Seeing how responses also relate to networking, scientific identity, and scientific literacy is of great interest. Continuing to apply CUREs to our biology laboratories will allow extrapolation of our results to a broader audience. Our experimental approach was

**Table 1.** Online survey responses comparing students from CURE and non-CURE courses. For Questions 1 and 2, only Yes and No were options.

<table>
<thead>
<tr>
<th>Question</th>
<th>CURE</th>
<th>Non-CURE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>1</td>
<td>88%</td>
<td>13%</td>
</tr>
<tr>
<td>2</td>
<td>100%</td>
<td>0%</td>
</tr>
<tr>
<td>3</td>
<td>38%</td>
<td>0%</td>
</tr>
<tr>
<td>4</td>
<td>100%</td>
<td>0%</td>
</tr>
<tr>
<td>5</td>
<td>38%</td>
<td>0%</td>
</tr>
</tbody>
</table>

**Figure 2.** Boxplot of TOSLS scores among the three different course sections, comparing the CURE and two non-CURE classrooms. Solid lines represent median values, boxes are upper and lower quartiles, whiskers show 75% and 25% quartiles, and dots mark outliers.
limited by the administration of TOSLS and PITS at midsemester and the end of the semester rather than at the beginning and end of the semester. Moreover, the online survey was administered only at the end of the semester. Our discoveries are specific toward a certain population of students.

Driving forward both ongoing and previous conclusions, our results garner confidence in the effectiveness of CUREs. We specifically add a great deal of confidence on impactfulness of CUREs in a biology laboratory environment. Furthermore, the data suggests an importance of quality research in piquing students’ interest in biological concepts.

Acknowledgments

We would like to thank A. Buchanan and A. Powell for valuable input. We would also like to thank Emporia State University for support and funding.

References


