Comparing the Outcomes of "pre-CURE" Compared to Inquiry-based Introductory Biology Labs

Allison L. Lansverk[1], Deborah A. Lichti[1,2], Kate W. Blinka[1], Kristine L. Callis-Duehl[1,3]

[1]Department of Biology, East Carolina University, Greenville, NC, 27858

[2]Interdisciplinary Science Learning Laboratories, University of Delaware, Newark, DE, 19716

[3]Education Research and Outreach, Donald Danforth Plant Science Center, St. Louis, MO 63017

Abstract

Course-based undergraduate research experiences (CUREs) are increasing being taught to undergraduates across institutions and topics and offer an opportunity for more students to participate in authentic research. There are many types of "CUREs" with varying degrees of including all five elements identified by Auchincloss et al. (2014) as defining a CURE course. Even when authentic research labs do not include all the elements of a CURE, they still provide students an opportunity to engage in the process of science and could be an important transition from inquiry-style labs to full CURE labs. In this paper, we call labs that provide research experiences, yet may include lower-stakes, non-iterative or narrowly relevant results "pre-CURE" labs. We sought to assess if students were developing a better understanding of experimental design and scientific argumentation in a pre-CURE vs. a guided-inquiry version of the same introductory biology lab, as evidenced by scientific paper writing assignments. Students in guided-inquiry labs performed better on argumentation themes (p=0.01549) in their lab reports, conceivably because they had narrower variable selection. However, students learned more about experimental design in the pre-CURE course (p=1.77^{-12}). Particularly because the pre-CURE gave the students the opportunity to fail by allowing them to design experiments that did not yield results.

Keywords: pre-CURE, CURE, citizen science, productive failure, laboratory, undergraduate

Introduction

Research experiences are an integral component of education for biology students. Experience conducting research has a major impact on the development of analytical and applied skills and the understanding of the relevant concepts, as well as a long-term impact on the student's continued passion for and involvement in science (Latzer et al., 2015). It has long been recognized that undergraduate research experiences (UREs) are an important part of training each subsequent generation of scientists (Kinkead, 2012). Apprenticeship has been the classical model for undergraduate research, and positive outcomes are reported for undergraduates who are able to participate in some sort of research endeavor (Gentile, Brenner, and Stephens, 2017). In 2009, the American Association for the Advancement (AAAS) of Science published "Vision and Change," a call to action to engage undergraduates in authentic science experiences (Woodin et al., 2010). One of the ways that students can participate in research is a course-based undergraduate research experience (CURE) (Brownell and Kloser, 2015).

Dolan (2016) defines CUREs as "learning experiences in which whole classes of students address a research question or problem with unknown outcomes or solutions that are of interest to external stakeholders," which is in contrast to more traditional apprenticeship-type research experiences that are often more or less one on one. This also differs from an inquiry-based model for a laboratory class since to be a true CURE, as with any scientific research, there is an unknown experimental outcome to the students and scientific community alike (Latzer et al., 2015, Dolan, 2015). CUREs are gaining use (Figure 1). Approximately 307 institutions of higher education, 17 of which are international, are currently registered within CUREnet, a network of programs and people creating CUREs established in 2012 by Erin Dolan (University of Georgia), Dave Micklos (Cold Spring Harbor Laboratory), and Nancy Trautmann (Cornell Lab of Ornithology). CUREnet estimates that at least one third of those are still actively teaching CUREs (*CUREnet*).

One reason why CUREs have become more prevalent lies in the numerous advantages of utilizing CUREs rather than more traditional UREs, such as internships. Many institutions do not possess the resources to involve large numbers of undergraduates in research in an internship capacity (Auchincloss et al., 2014). To circumvent this issue, CUREs have arisen as a popular alternative to apprenticeship

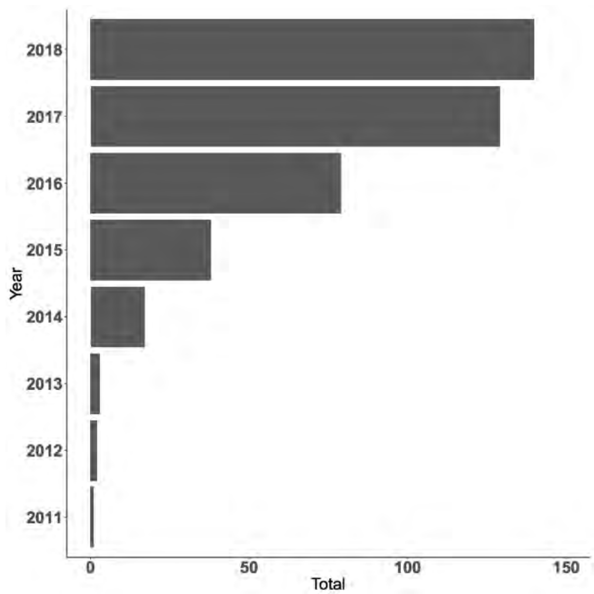

Figure 1. January 2019 Google Scholar advanced search of exact phrase "CUREs" AND "course-based undergraduate research experience" shows a steady increase of hits from the year 2011 until 2018.

research experiences (Bangera and Brownell, 2014). Another advantage to CUREs when compared to the classes they are often replacing (either cookbook labs or directed inquiry-based labs) is that they still engage students in the process of real science as an internship or apprenticeship would while addressing some of the limitations inherent in the latter. Students formulate their own research questions and create their own research designs, which can either result in success or failure. There are unknown outcomes to these research questions, and the results that the students obtain are genuine data that can support their hypotheses (Auchincloss et al., 2014).

However, implementing CUREs is not without its disadvantages and challenges. If students are unprepared for research, they cannot take full advantage of the CURE or URE. This can happen if they are not ready for the empirical process of science and experimental design. In an innovative study on student perceptions of research in 2005, Meyer, Shanahan, and Laugksch found that undergraduate students have problems with a general conceptual understanding of research even after taking standard biology classes. Murtonen (2015) argues that unless students are well-versed in the methodology in their discipline, they will not be fully successful. Additionally, students coming out of cookbook or inquiry-based labs are often not ready for the iterative process of a true CURE that involves potentially failing in their research process (Brownell and Kloser, 2015) and potentially needing to adjust either their research or methodology. Not all in-lab authentic research projects meet the five criteria laid out by Auchincloss et al (2014)—use of scientific practices, discovery, broadly relevant work, collaboration and iteration—yet these research projects still provide students valuable training in authentic research.

These authentic research courses that are not quite CUREs provide students with those foundational skills and understanding necessary to be successful in a full CURE or URE and we call these course-based authentic research projects a pre-CURE. In the context of this discussion, a pre-CURE is defined as a course-based undergraduate research experience that offers the opportunity to participate in the process of authentic science, but does not meet all the criteria of a CURE and/or at a level that is slightly more introductory than a true CURE. These laboratory courses provide authentic research experiences but may be missing iteration and/or the results may not be broadly relevant or applicable to the greater scientific knowledge or community. In other words, the data collected may not be intended for publication or greater communication. The lack of iteration or the limited scope of the results could be due to the limited time available in class for the project or the limited resources available to the course/school. Pre-CURE courses allow students to transition from open inquiry experiments, where the students develop an experiment for a question that has a known answer, to a full CURE, where students spend the course collecting data on a specific research question where the answer is unknown to science.

We expect that the advantages of a pre-CURE will be that the students will walk away feeling more comfortable with engaging in the process of science both through designing and troubleshooting, without the stress of collecting data for a high-stakes experiment. This in turn will allow them to engage in more complex and substantive experimentation in true CURE curriculum models or in UREs without the need to waste valuable time and resources implementing foundational basics. We will examine the advantage of an authentic research experience that does not meet all the CURE criteria, a pre-CURE, compared to an inquiry-based lab, on student's ability to collect data, analyze data, and troubleshoot problems.

For our pre-CURE, we chose to use a citizen science project as the framework for an authentic research context. Citizen science, which is subsumed under the umbrella of the field of public participation in scientific research (PPSR), engages the public in the process of scientific investigation through one or more steps of the following series of steps: asking a question, collecting data, and interpreting results.

Research that involves citizen science often necessitates the collection of data over a large geographic area or over a long period of time. These projects vary in scale. Some involve only few volunteers, while other span continents. These projects often have the dual objectives of scientific advancement and education. Though contributions of non-scientists to scientific discovery dates back to the time of Galileo or even much earlier, the earliest citizen science projects as they are currently defined date back to the 1800s (Cornell Lab of Ornithology). This sort of democratized approach to involving large numbers of laypeople in the expansion of scientific research represented a new trend in how scientific researchers approached and understood the principles of research outside of the traditional models.

For our independent pre-CURE, we made use of a citizen science project named Budburst. Originally known as "Project BudBurst," Budburst was developed in 2007 to connect citizen scientists with researchers, educators, and horticulturists to investigate plant life cycle events to help better understand impacts humans are having on the environment. We chose this project for the pre-CURE because it was free to participate, did not require any equipment, allowed for data collection at any time of the year, has been implemented in all 50 states by tens of thousands of people (*Budburst*), and, due to the ready availability of trees on campus, allowed students easy access to the research subject. The students designed their own questions regarding the study of leaf senescence or budding: for example, "how does temperature affect budding in flowering dogwood?" They then collected their own data from trees on campus and gathered previously documented data from the Budburst database.

The purpose of this study was to compare learning outcomes between the guided-inquiry version of an introductory biology lab course for majors at East Carolina University and a pre-CURE version of the same course. In this study, we sought to assess if students were developing a better understanding of experimental design and scientific argumentation in the guided inquiry versus the pre-CURE version of the lab. We investigated this by scoring research papers from both versions of the lab for a list of useful target skills with a validated empirical and representational skills rubric for argumentation.

Methods

Prior to the study, we obtained IRB approval and student consent (UMCIRB16-001669 and UMCIRB15-001990). Both a pre-CURE and the guided-inquiry were conducted in second-semester introductory biology labs during the Fall 2016 and an additional pre-CURE was conducted a second time during the Spring of 2017. The instructors of all of the sections, inquiry and pre-CURE, were the same, minimizing the risk of inconsistency posed by such variables as instructor or implementation influences by maintaining continuity. The same instructor also taught all the iterations of the pre-CURE lab. The learning objectives for all the labs was the same and covered both ecology and evolutionary biology themes. Note that the guided-inquiry-activity was geared towards the topic of evolution while the pre-CURE activity focused on ecology; both topics were taught in all second-semester introductory biology lab course.

We began our analyses by assembling the student papers from the guided-inquiry sections (henceforth referred to as the "inquiry lab") and the pre-CURE version of the lab (henceforth referred to as the "pre-CURE lab"). The student paper for the inquiry lab was a write-up about a *SimBio* virtual lab, a virtual simulation where students could manipulate environmental and population parameters to test natural selection (*SimBio*). The final paper for the pre-CURE lab was a paper on the experiment where students designed a research project collecting and using citizen science data from Budburst about leaf senescence (during fall semester) or leaf/flower budding (during spring semester).

Virtual Modeling Activity

Students in the guided-inquiry lab completed a virtual lab activity called Darwinian Snails outside of class in the second week of the semester using the program *SimBio*. The purpose of the Darwinian Snails activity is to help students investigate assumptions behind natural selection. The lab leads the students through simulations within the experimental system of crabs (predator) and snails (prey) following the work of Dr. Robin Seeley on the Gulf of Maine. Students come up with a hypothesis around competition and population density of snails and crabs, then manipulate a set of virtual variables to test their hypothesis. The students can adapt the parameters of the experiment by violating the assumptions of natural selection and investigate if evolution via natural selection still occurs. The virtual lab is broken into six exercises and the students are led through which parameters to change and then are provided with a graph to show the outcome of the experiment. Although each exercise has directions to follow while conducting the computer simulations, the students have the freedom to alter parameters and generate results, giving them the flexibility to run an experiment as well if they desire; this is why we are calling it "guided inquiry" or "inquiry-based". At the conclusion of the experiment, students were assigned a final paper to write up the results from the virtual lab. The instructors then broke the papers down into sections relating to typical divisions found in scientific papers (methodology, results, discussion, etc.). The students were given back each draft with edits and comments, which they then had to revise into one edited final paper.

Budburst Project

The pre-CURE sections used citizen science data from Budburst. Students, in groups of four, designed a project that required data collection from the field and the pulled from the Budburst database. Each group developed a question based on tree phenology (i.e. first bud or leaf senescence) for Greenville, NC, and another location throughout the country, incorporating one abiotic variable (e.g. temperature, rain fall). Students then collected local data and pulled data from databases that met their research question needs. The data were analyzed, and graphs were developed using JMP during the Fall 2016 semester and the provided R markdown files in RStudio during Spring 2017 semester. Finally, each member of the group presented their findings with an individual scientific paper. We designed the module hoping to provide the opportunity for higher-order thinking with the students learning how actual research occurs and experiencing the trials and tribulations of working in groups, and collecting field- and web-based data.

Student demographics

The majority of the students who take this laboratory course and who were included in this study are freshman and sophomores who are STEM majors. It is important to note that while most students are required to have first-semester introductory biology as a pre-requisite to the pre-CURE lab and inquiry lab, exceptions occur and therefore we cannot not guarantee that all students involved in the study had one semester of biology lab prior to this study. The first semester introductory biology course is a guided-inquiry style lab course focused on cell and molecular biology topics and does have one open-inquiry experiment. Therefore, most students entering this course have had the opportunity to create a question, hypothesis, and design an experiment prior to the course.

Study design

We compiled all of the final papers from the three sections and included all papers that were complete (from abstract through discussion sections) for our analysis. This resulted in 19 papers from the Fall 2016 inquiry lab, 18 papers from the Fall 2016 pre-CURE lab, and 18 papers from the Spring 2017 pre-CURE lab.

Using Excel, we randomly assigned each paper an anonymous student identification number and all identifying information about the student was removed. Each paper was then assigned to two out of four potential reviewers using a random number generator in Excel. Two of the reviewers were previous instructors of the course, but identifying information was removed and papers were randomized to minimize bias. The other two were non-instructors and were included to further decrease bias since pairs of reviewers had to come to a consensus on scoring. Reviewers used an adapted and validated rubric based on the Assessment of Scientific Argumentation in the Classroom (ASAC) to evaluate laboratory argumentation skills (Walker et al., 2018, Walker et al., 2019). The rubric was comprised by of 23 target skills that we believed would be useful and were reasonable for science majors to obtain during an introductory biology course (Supplemental Table 1). These skills were determined prior to the start of the study and agreed upon by all the researchers involved as well as external researchers who provided guidance on the development of the rubric. The skills rubric reflects the "ideal" learning objectives that we would hope to see demonstrated by second-semester biology majors.

For each target skill, the rubric asked a question about whether the student demonstrated proficiency of the skill in question. The rubric was entered into Qualtrics (Qualtrics, Provo, UT) and set up as a survey that could be completed. We re-validated the adaptations of the rubric by using it to code randomly selected lab reports and then discussed additions or modifications to the rubric necessary to capture the skills in the lab reports. After we were confident in the rubric, all four reviewers scored the subset of papers assigned to them.

For each question in the rubric, we assessed whether every student demonstrated an acquisition of each target skill in their final paper and scored accordingly in Qualtrics. Once all the results were tabulated, we downloaded the data and compared the results between each pair of reviewers for each student's paper. We identified any discrepancies in the pair of reviewers' answers for each question for every student paper. After identifying the discrepancies, the first step was to verify that we did not simply miss something during the initial scoring process. To do this, reviewers went back through their answers and looked for the examples that the other reviewer had identified to see if they had been overlooked. For those cases where discrepancies still remained, the reviewers met as pairs to discuss the thought process for each set of discrepancies. Reviewers reached 100% inter-rater agreement in cases when assessment was not initially in agreement. In cases where the two reviewers could not reach agreement, a third reviewer was employed to resolve the impasse.

Once the data was reconciled, it was imported into RStudio for analysis. Because the data was not distributed normally (as confirmed by a Shapiro test), we used a Mann-Whitney U Test to assess if there were differences in learning outcomes for the 23 target skills between the students in the inquiry lab and the students in the pre-CURE labs. The data for the two sections of the pre-CURE labs were also compared to each other using the same methods.

Results

Of the 23 target skills, 13 showed a significant difference between the inquiry and pre-CURE labs. We identified four of these target skills as important for future CURE/URE projects. These four skills include: constructing a scientific question, designing an experiment, representing data in a student generated form, and supporting the claim with evidence. Due to their importance in future research success, we discuss these results in greater detail below.

Constructing a scientific question was 83% higher for students in the pre-CURE course compared to the inquiry course (p=0.009037) (Figure 2). Designing an experiment was 100% higher for students in the pre-CURE course compared to the inquiry course (p=1.77^{-12}) (Figure 2). Representing data in a student generated form was 100% higher for students in the pre-CURE course compared to the inquiry course (p=1.25E-09). Supporting the claim with evidence was 43% lower for students in the pre-CURE course compared to the inquiry course p=0.01549) (Figure 2).

In addition to the four key factors, nine other elements of interest were significantly different in pre-CURE courses compared to inquiry courses, locating information relevant to a scientific problem (p=4.92E-05); designing an experiment that appropriately answered the question (p=0.0123); appraising an experiment design; creating an effective figure label (p=0.001683); interpreting visual representations of data (p=5.66E-06); justifying the claim (p=0.002661); identifying additional information needed to support an argument (p=0.02457); providing alternative explanations(p-value 0.000969); and using appropriate scientific language (p=0.0257).

When comparing the two iterations of the pre-CURE, we found significant differences for 3 of the 23 target skills: locating information relevant to a scientific problem; designing an experiment that appropriately explores the question; and using appropriate scientific language. We found no significant difference between the two iterations and therefore we combined the data from the two iterations.

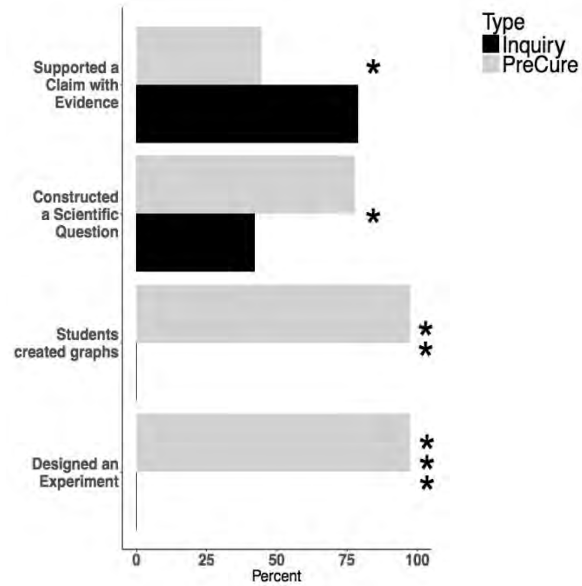

Figure 2. Percentage of assessed students that accurately met target skill for the Inquiry lab and pre-CURE lab. The four skills are: 1) Supported a claim with evidence, 2) Constructed a scientific question, 3) Students created graphs, 4) Designed an experiment. Number of asterisks reflect level of significance (more asterisks = smaller p-value).

In a follow-up three years after the course, we found that 26.33% of students who took the inquiry-style lab section changed their major from a STEM major to a non-STEM major while only 16.7% of students in the pre-CURE lab did the same. More students from the pre-CURE sections also went on to take CURE labs, 16.7% in the pre-CURE sections versus 10.5% in the inquiry-style lab. However, 31.6% of students in the inquiry-lab section were more likely to engage in a traditional undergraduate research experience while only 22.2% of students in the pre-CURE lab did the same.

Discussion

It is our conclusion that citizen science projects provide appropriately scaffolded frameworks for pre-CURE projects for a number of reasons. First, the data are easy to access since they are aggregated in one online database, the utility of which was demonstrated by Budburst. Second, the data are relatively easy to collect, with detailed protocols that are written in a way that allows non-experts to faithfully execute data collection. The advantage of a citizen science project is that it can involve non-scientists, so the level of skill that is required to collect the data is easily accomplished by intro-level biology undergraduate students. These projects offer students the ability to contribute to real science, with low barriers to participation: in the case of Budburst specifically, for example, there is no expensive equipment involved or the need for costly analyses. Finally, the plethora of citizen science projects available allows instructors to choose projects with research questions related to areas of a student's life that are neither abstract nor nebulous (e.g. leaf color change on trees) and fit with the learning objectives of a wide variety of course content.

In this study we sought to investigate if participating in a pre-CURE contributes to an increase in the skills we identified as necessary to be successful in undergraduate research experiences and future upper-level CURE courses. We found that for over half of the target skills that we assessed, there were significant differences between the students who participated in the pre-CURE and those in a traditional inquiry class. Three of these target skills (Figure 2) we see as being very important for students to develop prior to taking a CURE, especially with regards to understanding methodology, the importance of which was highlighted by Murtonen (2015). These target skills are constructing a scientific question, designing an experiment, and representing data in a student generated form. Therefore, it appears that pre-CURE courses may provide students with a greater foundation of skills applicable to science research than the skills taught in traditional inquiry labs.

In the case of designing an experiment and representing data in a student-generated form, the important aspect to note is that the pre-CURE lab presented the students with the opportunity to participate in these steps in a way that the guided inquiry class did not. In the case of graph generation, in the pre-CURE course, the students entered the data into the graphical software of their choice (excel, R, or JMP) and then decide what type of graph to generate to best fit their methodological choices; in contrast, for the inquiry lab, the computer program would automatically generate the relevant graphs at the end of the sections after the students had followed the directions for the simulations. As for experimental design, while the students in the inquiry lab did develop hypotheses and test those hypotheses, they were not asked to design their own experiment; their independence was restricted to selecting a few variables and adjusting the levels of the selected variables, with the virtual lab itself generating the graphs once the settings were selected. Because the pre-CURE lab offered the opportunity to be fully immersed in all steps of the methods, we see this as a beneficial educational experience for these students going forward (Murtonen, 2015).

The last target skill that differed between the two lab treatments demonstrated the ability to support the claim with evidence. While both sets of students struggled with argumentation and science communication, supporting a claim with evidence was done less effectively (43% lower) by students in the pre-CURE course (p=0.01549; see Figure 2). In the pre-CURE course, there were a wider range of ecological variables to select from when the students designed their experiment. They could choose any variable they wanted that might affect leaf senescence in the fall/budding in the spring: e.g. amount of sunlight, level of rainfall, etc. In addition to this, nature itself was manipulating these variables; this applied aspect to the experiment led to more diverse outcomes compared to a computer simulation. The less effective examples of students supporting claims with evidence could be due to the challenge of interpreting a broader selection of variables compared to the restrictions imposed by the guided inquiry course. What this means is that the open nature of the pre-CURE course gives more opportunity to fail, but by the same token, it also provides a chance to learn from the failures. Although at first this might sound undesirable, the concept of "desirable difficulties" has seen support in the fields of psychology and mathematics (Bjork, 1994a, 1994b, Kapur and Bielaczyc, 2012). Bjork (1994a, 1994b) found that long-term retention increased when difficulties were introduced because these forced the students to think more closely, and extensively, about the processes involved. Productive (or constructive) failure has also been shown to assist students in problem solving when compared to students who received direct instruction (Kapur and Bielaczyc, 2012). There has been little investigation thus far on the impacts of productive failure in a biology laboratory setting; however, pre-CURE courses could offer researchers the opportunity to investigate this since the nature of a pre-CURE course gives students the opportunities to reason through why a result did not come as expected and what factors might be at work. Similarly, productive failure helps students understand iterative nature of science so when participating in other research projects they would hopefully be better able to cope with the setbacks that come with research.

We saw in increase in the percent of students from the pre-CURE lab course taking CURE labs in the subsequent three years, while inquiry-lab students were more likely to engage in a more traditional undergraduate research experience in the laboratory of a professor. We suspect that these differences are due to the pre-CURE students being aware of the possibility of engaging in authentic research through CURE courses, while inquiry students were not as familiar with this pathway and therefore followed a more traditional research path. However, further follow-up with the students will be needed to determine if this is the underlying cause of the difference in research pathways.

The definition of a CURE is still being discussed within the community of Biology Education Researchers (Corwin et al 2018), but as we decide to what extent all criteria laid out by Auchincloss et al (2015) are necessary for undergraduates to receive a learning benefit from conducting authentic research in laboratory classes, determining the benefit of introducing even limited versions of authentic research, pre-CUREs, can assist programs in making difficult choices around laboratory course reform. From our data, we feel that because of the benefits in helping the students to learn vital scientific processes and communication, and to learn from failures prior to taking a full CURE or participating in research, we recommend implementing pre-CURE modules and courses in early introductory science labs. Therefore, despite results showing students need >36 hrs. of engagement in research to benefit fully from the experience (Shaffer et al. 2014), we found that even a more limited authentic research project within a laboratory class is beneficial to students preparation as scientists.

Acknowledgements: We would like to thank Jonathan Hall, Joi Walker, and Durin Booth for their contributions making this research and manuscript a success.

Funding: This work was internally funded.

References:

Auchincloss, L. C., Laursen, S. L., Branchaw, J. L., Eagan, K., Graham, M., Hanauer, D. I., & Towns, M. (2014). Assessment of course-based undergraduate research experiences: a meeting report. CBE—Life Sciences Education. *13*(1): 29-40.

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.

Bjork, R. A. (1994). Institutional impediments to effective training. Learning, remembering, believing: Enhancing human performance. 295-306.

Bjork, R. A. (1994). Memory and metamemory considerations in the. Metacognition: Knowing about knowing. 185.

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.
DOI: 10.1080/03075079.2015.1004234

Budburst. Retrieved November 10, 2018 from https://budburst.org/

Cornell Lab of Ornithology. (accessed January 15, 2019). Retrieved from http://www.birds.cornell.edu/citscitoolkit/about/history

Corwin, L A., E.L. Dolan, M.J. Graham, D.I. Hanauer, N. Pelaez. The Need to Be Sure About CUREs: Discovery and Relevance as Critical Elements of CUREs for Nonmajors. Journal of Microbiology & Biology Education. 19(3): 1-5.

CUREnet Course-based Undergraduate Research Experience. (2018, October 11) Retrieved November 10, 2018 from https://serc.carleton.edu/curenet/index.html

Dolan, E. L. (2016). Course-based undergraduate research experiences: Current knowledge and future directions. Washington, DC: National Research Council.

Kinkead, J. (2012). What's in a name? A brief history of undergraduate research. CUR Quarterly. 33(1): 20-29.

Kapur, M., & Bielaczyc, K. (2012) Designing for Productive Failure, Journal of the Learning Sciences. 21(1) 45-83. DOI: 10.1080/10508406.2011.591717

Latzer, J., Mans, T., & Provost, J. (2015). What is a CURE? In A practical guide to course-based undergraduate research experiences.

Meyer, J. H., Shanahan, M. P., & Laugksch, R. C. (2005). Students' Conceptions of Research. I: A qualitative and quantitative analysis. Scandinavian journal of educational research. 49(3): 225-244.

Murtonen, M. (2015). University students' understanding of the concepts empirical, theoretical, qualitative and quantitative research. Teaching in Higher Education. 20(7): 684-698.

Gentile, J., Brenner, K., & Stephens, A. (Eds.). (2017). Undergraduate research experiences for STEM students: Successes, challenges, and opportunities. National Academies Press.

Qualtrics. Retrieved November, 2018 from https//www.qualtrics.com

SimBio. Retrieved November 10, 2018 from http://simbio.com/

Shaffer, C.D., Alvarez, C.J., Bednarski, A.E., Dunbar, D., Goodman, A.L., Reinke, C., Rosenwald, A.G., Wolyniak, M.J., Bailey, C., Barnard, D., Bazinet, C., Beach, D.L., Bedard, J.E.J., Bhalla, S., Braverman, J., Burg, M., Chandrasekaran, V., Chung, H.-M., Clase, K., Dejong, R.J., Diangelo, J.R., Du, C., Eckdahl, T.T., Eisler, H., Emerson, J.A., Frary, A., Frohlich, D., Gosser, Y., Govind, S., Haberman, A., Hark, A.T., Hauser, C., Hoogewerf, A., Hoopes, L.L.M., Howell, C.E., Johnson, D., Jones, C.J., Kadlec, L., Kaehler, M., Silver Key, S.C., Kleinschmit, A., Kokan, N.P., Kopp, O., Kuleck, G., Leatherman, J., Lopilato, J., Mackinnon, C., Martinez-Cruzado, J.C., McNeil, G., Mel, S., Mistry, H., Nagengast, A., Overvoorde, P., Paetkau, D.W., Parrish, S., Peterson, C.N., Preuss, M., Reed, L.K., Revie, D., Robic, S., Roecklein-Canfield, J., Rubin, M.R., Saville, K., Schroeder, S., Sharif, K., Shaw, M., Skuse, G., Smith, C.D., Smith, M.A., Smith, S.T., Spana, E., Spratt, M., Sreenivasan, A., Stamm, J., Szauter, P., Thompson, J.S., Wawersik, M., Youngblom, J., Zhou, L., Mardis, E.R., Buhler, J., Leung, W., Lopatto, D., Elgin, S.C.R., 2014. A course-based research experience: how benefits change with increased investment in instructional time. CBE Life Sci. Educ. 13: 111–130.
https://doi.org/10.1187/cbe-13-08-0152

Walker, J., Van Duzor, A., & Lower, M. (2018). Facilitating productive argumentation in the laboratory. In Abstracts Of Papers Of The American Chemical Society (Vol. 255). 1155 16th St, Nw, Washington, Dc 20036 Usa: Amer Chemical Soc.

Walker, J. P., Van Duzor, A. G., & Lower, M. A. (2019). Facilitating Argumentation in the Laboratory: The Challenges of Claim Change and Justification by Theory. Journal of Chemical Education. 96(3):435-444.

Woodin, T., Carter, V. C., & Fletcher, L. (2010). Vision and change in biology undergraduate education, a call for action—initial responses. CBE—Life Sciences Education. 9(2): 71-73.

Supplemental Table 1. Target Skills used to assess written papers of pre-CURE lab and Inquiry lab students. Notes under good examples are quotes from papers that demonstrate a high score on that target skill.

Target Skills	Good Example from Inquiry	Good Example from PreCURE
Locate information relevant to a scientific problem. i.e. Identify databases to search for information in the field relevant to a problem. (Number of times)	[None utilized]	“. Budburst 2. Weather Underground history database
Construct a relevant/appropriate scientific question for a given problem. i.e. Develop a question that addresses a key issue in a given scientific problem or scenario.	In essence the lab experiment will be testing the evolution of snail by natural selection.	Does [the amount of] rain fall have an effect on the blooming stages of Prunus Serrulata (Japanese Cherry Blossom)?
Generate a hypothesis or make a prediction based on a scientific model. i.e. Identify an appropriate scientific model and apply it to the given problem in developing a hypothesis or making a prediction.	The hypothesis of the lab is that the introduction of the Green European Crabs did impact the shell thickness of the snails as the thinner shells were easily cracked by the crabs thus their population decreased.	Our hypothesis for this question was if the annual temperature continues to increase, we expect to see more green foliage present during autumn.
Design an experiment to test a scientific question. i.e. Develop a study that identifies relevant factors and collects data to effectively examine the problem.	[None created. Followed pre-determined steps]	Five post oaks were selected at various points around campus. Then, on four separate occasions throughout the duration of two weeks, the trees were visited. On the initial visit, coordinates were recorded along with the weather, temperature, and percent coloration for the day. Then pictures were taken of the leaves and overall tree. On the rest of the visits weather, temperature, and percent coloration were recorded in order to track the progression of color change present in the leaves. During each of the three remaining visits photos of the trees were taken in order to compare coloration changes to the previous visit.
Designed an experiment that appropriately answers question. i.e. Used appropriate techniques to answer question.	[None created. Followed pre-determined steps]	Repeated 5 times over 20 days to find the trend.
Appraise an experiment design to identify elements and limitations and how they impact scientific findings/conclusions. i.e. Identify strengths and weaknesses of research design elements (e.g., potential sources of error, variables, experimental controls). (Number of times)	[None created. Followed pre-determined steps]	To start out this experiment, three different species of trees are to be picked, and data will be collected from three trees of each species for a total of nine trees. This stabilizes the data, acting as insurance in case one of the trees is sick and presents false data
Troubleshoot technical issues. i.e. Evaluate a scientific problem to identify possible technical causes. (Number of times)	[None.]	Students had to address everything from hurricanes impacting their study trees to accessing the trees due to construction.
Represent data in a visual form (student generated). i.e. Convert relevant information into an appropriate visual representation given the type of data (count number of graphs/figures). (Number of times)	[None created. SimBio generated them automatically]	Three graphs

Created an effective figure/graph label. i.e. Label gives meaningful information about the data represented. (Number of times)	Figure 1. Flat periwinkle shell thickness after the introduction of European green crabs.	1. Figure 1: Mean percentage of leaves shed at a set date comparing the leaf fall in the mountain and coastal regions of North Carolina. 2. Figure 2: One-way analysis and t-test showing the difference between the mean values of leaves shed in the mountain region versus the coastal region of North Carolina.
Interpret visual representations of data. i.e. Interpret or explain information presented in visual forms. (Number of times)	After being regenerated again the snails still stayed the same showing the same results. (Fig 5.). The shells with the thickness of 6,7, and 8 were the ones left on exercise 1. The snails with less thick shells were mainly eaten because the crabs didn't have to use as much time to try and crack them open. Exercise 2 the shell thickness was generally the same this population never changed because there were no other crabs to reproduce due to having a selective amount of genes. Crabs ate the snails faster during this trial since there was no heritability amongst the snails. In the third exercise mutation would only allow the snails to reach a thickness no higher than 4 inches. Reproducing the snails with this mutation they generally became thick as a population.	"To analyze our data, we chose to create a scatterplot including a line of best fit that compared the temperature measured in Fahrenheit to the color of the leaf measured in wavelengths. The line showed a slope of 0.0115 and an r-squared value of 0.01742 "
Evaluate evidence.	The population of snails without any predatory factors has a random assortment of shell thicknesses. This makes since because no specific shell type is being selected for. The population of the snails being affected by predatory selective pressures, tend to have extreme shell thicknesses, since the crabs could easily devour thinner shelled snails.	The P value calculated with one-way ANOVA analysis showed that Sweet gums were the only species that had a p value below .05. Both the Pine and the Magnolias had p values higher than .05. This means we accept the null hypothesis for Pine and Southern Magnolia and reject the null for Sweetgum.
Statistics techniques utilized. (Chi sq., T test, Correlations, Averages, ANOVA, Other)	[None. The program provided all analyses and graphs without the student's input]	Calculated r squared
Stats techniques were used appropriately for data.	[Since there was no student choice in the analysis, there was no way for them to use inappropriate statistics]	When we looked at our scatter plot, the slope of the line was 0.0115 which if were rounded would be zero. The r-squared value of 0.01742 would also be zero if rounded.

Make a claim based on evidence i.e. Make a claim that answers a question or provides a solution to a problem.	After reviewing my results and gathered information, I noticed a change in the shell thickness overtime judging from the graphs. The shells in the earlier years were less thick as evolution changed them into thicker harder shells. Yes, this snail population over time meets the requirements for evolution through natural selection.	The P value calculated with one-way ANOVA analysis showed that Sweet gums were the only species that had a p value below .05. Both the Pine and the Magnolias had p values higher than .05. This means we accept the null hypothesis for Pine and Southern Magnolia and reject the null for Sweetgum.
Support the claim with evidence. i.e. Gives specific examples.	The assumed hypothesis was that the introduction of The Green European Crabs impacted the population of snails in that the thicker shelled snails had a greater level of fitness. This explanation is justified by the fact that the thin shelled snails' population only decreased upon their arrival.	The P value calculated with one-way ANOVA analysis showed that Sweet gums were the only species that had a p value below .05. Both the Pine and the Magnolias had p values higher than .05. This means we accept the null hypothesis for Pine and Southern Magnolia and reject the null for Sweetgum.
Justify the claim. i.e. Justify why the evidence is appropriate.	Having known about Darwin's three requirements for adaptive evolution, the results supporting natural selection by providing evidence of variance, increased fitness, and heritability were not surprising to me. In each simulation where one or more of the three requirements were removed no evolution appeared, which followed similar logic and was also expected.	While this might show that temperature plays an effect, more data points are needed to determine that. This is a problem that our group experienced in testing our hypothesis.
Identify additional information needed to support an argument. i.e. Explain how new information may contribute to the evaluation of a problem. (Number of times)	These flat periwinkle snails should be studied further to see how they evolve naturally over the coming years. Shells may not be the only thing about the snail populations that change. Traits such as the size of the snail or shell color may also change due to natural selection.	Our result supported our hypothesis but that doesn't mean there were not questionable factors throughout our experiment. For example, the amount of daylight increased by three minutes from day one of collection to day two but there was still evidence in all nine trees of color change. After looking back at our data collection sheet, we realized that the second day was mostly cloudy from the rain in the early morning. This means that the sun radiation was lower than that of the first day. This is most likely the reason for why the increase in percent change in leaf color was present. So, if this experiment were to be carried out for a second time, it may be helpful to take in consideration the amount of radiation the sun is giving off for the days of collection. Upon duplicating this experiment, it would also be helpful to stretch out the data collection process for another two or three weeks. In doing this experiment, there was not enough time to collect the amount of samples needed to give accurate data. Our chart may have been more consistent and possible even have a specific rate of change if we had the time to extend the process of data collection for each of the nine trees.

Provide alternative explanations for results that may have many causes. i.e. Recognize and explain possible alternative interpretations for given data or observations. (Number of times)	[Did not do]	One factor that could have influenced the variability of the data was the constant change of the temperature. It would be at one extreme during the beginning of the week and at the other extreme at the end of the week. Another factor that might have played a role on the strong variability was the aftermath effects of hurricane Matthew on the vegetation around campus. The hurricane brought strong winds and excessive flooding that could have had harmful effects on the trees. This could also explain why some of the trees were already missing a good amount of leaves. It made it hard to determine a precise estimation of coloration for the trees that did not have as many as the others. This could have also skewed the data some since there was not as many leaves to look at for coloration purposes. Early morning frost was another factor that could have affected the overall variability between the trees. Frost indicates cooler temperatures, which signals the trees to stop chlorophyll production.
Integrate and apply knowledge across subdisciplines. i.e. Identify and use concepts from previous courses to interpret given data and observations. (Number of times)	[Did not do]	[Did not do]
Uses appropriate Language. i.e. Uses scientific terms correctly.	[only examples if used if incorrectly]	[only examples if used if incorrectly]
References appropriate literature. i.e. Cites relevant primary literature. (Number of times)	Seeley, Robin Hadlock. 1986. Intense natural selection caused a rapid morphological transition in a living marine snail. Proceedings of the National Academy of Sciences, USA 83: 6897-6901. Trussell, Geoffrey C. 1996. Phenotypic plasticity in an intertidal snail: The role of a common crab predator. Evolution 50: 448-454.	1. Günthardt-Goerg, M. S., Kuster, T. M., Arend, M. and Vollenweider, P. (2013), "Foliage response of young central European oaks to air warming, drought and soil type." Plant Biology, 15: 185-197. 2. Lee, D. W., O'keefe, J., Holbrook, N. M., & Feild, T. S. (2003). "Pigment dynamics and autumn leaf senescence in a new England deciduous forest, eastern USA." Ecological Research, 18(6), 677-694. doi:http://dx.doi.org/10.1111/j.1440-1703.2003.00588.x

Critiques cited literature. (Number of times)	Seeley drilled a whole into different shell types and placed the snails into an ecosystem with a higher or lower level of invasive crabs. Over a few weeks, she examined that the snails with low spiral-shells had a higher survival rate against the European green crab. Seeley concluded that the crabs helped evolved the change in morphology of the snail, but it was natural selection that allowed these variations to occur in the population of snails (Seeley 6897).	Climate change has already affected the growth of some trees. In a study done by a team of scientists from the American Institute of Biological Sciences in 2007, 130 tree species, including some well-known trees such as the Loblolly pine, the sweetgum, and many types of oak, birch, and maple trees, were observed and their tolerance to climate change was determined. Using their data, the scientists have projected that by the end of the century, only one of the tree species that they studied will be able to grow in much of the southern United States (McKenney, Pedlar, Lawrence, Campbell, & Hutchinson, 2007). Although they did not study crepe myrtles directly, a lot of the species have similar growth conditions to crepe myrtles, so examining this data can still provide a prediction for how well they will withstand the projected climate change. This means that, in the next century, crepe myrtles may not be found in the southern United States.