Reports such as Vision and Change in Undergraduate Biology Education call for integration of course-based undergraduate research experiences (CUREs) into biology curricula and less emphasis on “cookbook” laboratories. CUREs, often characterized by a single open-ended research question, allow students to develop hypotheses, design experiments, and collaborate with peers. Conversely, “cookbook” labs incentivize task completion and have pre-determined experimental outcomes. While research comparing CUREs and “cookbook” labs is growing, there are fewer comparisons among CUREs. Here, we present a novel CURE built around an invasive grass, Bromus inermis. We evaluated this CURE’s effectiveness in improving students’ understanding of the Vision and Change competency relating to the application of the scientific process through development and testing of hypotheses. We did so by comparing changes in pre- and posttest scores on the Experimental Design Ability Test (EDAT) between Brome CURE students and students in a concurrent CURE, SEA-PHAGES. While students in both CUREs showed improvements at the end of the semester, Brome CURE students showed a greater increase in EDAT scores than did SEA-PHAGES CURE students. Additionally, Brome CURE students had significantly higher gains in 6 of the 10 EDAT criteria. We conclude that the Brome CURE is an effective ecological parallel to the SEA-PHAGES CURE and can help students gain a meaningful understanding of Vision and Change competencies.

INTRODUCTION

In 2011, the American Association for the Advancement of Science (AAAS) released a report titled Vision and Change in Undergraduate Biology Education: A Call to Action, describing the need for a shift from curricula reliant on memorization of facts toward curricula that incorporate core concepts and competencies that cut across biological scales and disciplines (1). Vision and Change described these core concepts and competencies as “the distinguishing features of undergraduate biology education, providing a strong foundation to guide the development of curricular frameworks.” The core competencies were meant to engage students in “real biology,” and the report encouraged educational institutions to integrate these competencies into their curricula.

One way in which Vision and Change competencies have been incorporated into undergraduate classrooms is through course-based undergraduate research experiences (CUREs) (2–6). These are often characterized by a single research question that is examined throughout a course or semester, open-ended outcomes, student-designed experiments, and peer-to-peer collaboration (7). Furthermore, CUREs expose students to much of the scientific process, including making observations, question/hypothesis development, experimental design, as well as data collection and data analysis/interpretation (8, 9). Ideally, a CURE is structured around a broadly relevant scientific topic/problem and allows students to produce novel insight into that topic (7). CUREs stand in sharp contrast to the traditional practice of “cookbook” laboratory exercises (discouraged by Vision and Change), where outcomes are often pre-determined and students are often intellectually disengaged because task completion rather than student curiosity is incentivized (10).

Despite a growing call for the integration of CUREs into undergraduate biology curricula and an increase in the number of CUREs, adoption of CUREs has been relatively slow. One challenge to CURE implementation may be the difficulty in assessing the impact of CUREs (7). Studies of
CURE effectiveness have utilized self-reported student perceptions of learning gains—often including a non-random set of students who enroll in the CURE voluntarily—that are sometimes examined without an appropriate comparison group, all of which introduces bias into any assessment of the CURE (7, 11). Summative course-based assessments such as professional conference-style presentations or journal article–style final papers may reflect a deep understanding of a particular CURE (4) but may be limited in their ability to assess whether the student can transfer the framework from that CURE to another, unrelated experimental context (12, 13).

Here we present a general overview and evaluation of a novel CURE designed to help students grasp the key competency relating to the application of the scientific process through the development and testing of hypotheses (1). This novel CURE (hereafter Brome CURE) is built around controlling the spread of an invasive grass species, Bromus inermis (smooth brome grass). To assess the effectiveness of the Brome CURE in addressing this Vision and Change competency, we compared first-year biology students enrolled in the Brome CURE with first-year biology students enrolled in a concurrent CURE, SEA-PHAGES, sponsored by the Howard Hughes Medical Institute (HHMI). Like the Brome CURE, the SEA-PHAGES CURE aims to “introduce students to research methods and approaches, experimental design, and data interpretation” (14). It has been implemented at a diverse range of institutions (15) and its positive impact on undergraduate persistence in STEM and self-efficacy is established (14, 16).

Having a single department, with all faculty committed to incorporating Vision and Change principles in the classroom, run multiple sections of different CUREs concurrently presented a rare opportunity. The type of comparison this allowed is unique: the few studies examining the effectiveness of CUREs either lacked a concurrent comparison group or used a concurrent “cookbook” experience as the comparison group (5, 11). While studies comparing CURE and “cookbook” systems highlight the positive impacts that CUREs can have on student learning and are important to our understanding of student learning (5, 11), assessing the effectiveness of a new CURE in comparison with a well-established CURE allows for an “apples-to-apples” comparison.

Specifically, we used the Experimental Design Abilities Test (EDAT) (13) in our evaluation. We administered the test to students in both CUREs at the beginning and end of the semester and examined changes in the scores. The EDAT, as the name indicates, allows students to demonstrate their ability to design an experiment in response to a “claim” statement. The EDAT is a content-independent assessment administered in a pre/post format, at the beginning and end of a course. This assessment tool is appropriate to evaluate CUREs because hypothesis evaluation and experimental design are central features of CUREs (4–7). Furthermore, the EDAT has been used in both major and nonmajor courses to assess the experimental design skills of introductory biology students (17).

**METHODS**

During the course of a single semester, Brome CURE students were asked to carry out self-designed experiments comparing the response of the invasive grass Bromus inermis and a noninvasive grass, Lolium perenne (as a control species), to treatments chosen by the students. Students in the Brome CURE collected experimental data using standard ecological metrics (e.g., biomass, above- and belowground allocation). By examining the invasion of this grass, students gain familiarity and competency with critical scientific skills, such as hypothesis development and testing, experimental design, and data interpretation, while also investigating a biological problem that has “broader relevance that extends beyond the course” (7). Concurrently, students in the first semester of the SEA-PHAGES CURE investigated the distribution of bacteriophage extracted from soil samples taken from various student-chosen sites on campus (14). Students in the first semester of the SEA-PHAGES CURE then carried out a series of laboratory techniques to isolate and identify the novel bacteriophage (including aseptic technique, streak plating, DNA extraction, and titer analysis). To be clear, this work is examining the impact of the Brome CURE system (a one-semester CURE) relative to the first semester of the SEA-PHAGES CURE, which can be taught as a two-semester sequence (14). The positive impact of the SEA-PHAGES CURE on undergraduate persistence in STEM and self-efficacy is established (14, 17). The comparison between the two CUREs is merely meant to provide a point of reference for the Brome CURE with respect to a nationally recognized CURE that also addresses goals outlined in Vision and Change (14).

To evaluate the effectiveness of the Brome CURE in addressing students’ grasp of the selected Vision and Change competency, we collected EDAT data from first-year biology students in one of two introductory biology courses featuring different CUREs, the Brome CURE and SEA-PHAGES CURE. The data were collected between the fall 2013 and fall 2015 semesters. Students in this first-year course were randomly enrolled in either an introductory section built around the Brome CURE system or the first semester of the HHMI SEA-PHAGES system. The students did not have an option of volunteering for a CURE vs. traditional “cookbook” structure, which helps avoid a documented bias in the assessment of CUREs (11). Across all years, a total of 160 students participated in the Brome CURE and 123 students participated in the SEA-PHAGES CURE. There were 18 sections across all years, with 10 sections of the Brome CURE system and 8 sections of the SEA-PHAGES system. Although the syllabi differed between the two CUREs, in any given year all the Brome CURE sections followed the same syllabus and all the SEA-PHAGES CURE sections followed the same syllabus. While there were minor changes to each CURE across years (e.g., exact timing of when statistics was taught during the semester), the major components of each CURE did...
not change and the EDAT was administered in the same manner across all years and sections.

In both CUREs, students were given the EDAT prompt on the first day of the course with minimal instruction, and the students were assessed again with a similar prompt at the end of the course (13). As recommended by Sirum and Humburg (13), the pretest EDAT scores and the EDAT scoring rubric were not shared with students during the semester. The post-test was administered in class at the end of the semester, before the final exam. Students were given a small amount of course credit for completing each test (pre and post).

The EDAT prompt invites students to describe an investigational design based on a claim about a health supplement. This prompt does not require the students to have any prior knowledge of the claim and is independent of the content in both CUREs. Both pre- and post-tests were scored by members of the Doane University biology department using the 10 scoring criteria described by Sirum and Humburg (13):

1. Recognition that an experiment can be done to test the claim.
2. Identification of what variable is manipulated.
3. Identification of what variable is measured.
4. Description of how the dependent variable is measured.
5. Understanding of the placebo effect.
6. Realization that there is one other variable that must be held constant (versus no mention).
7. Realization that there are many variables that must be held constant (versus only one or no mention).
8. Understanding that the larger the sample size or number of subjects, the better the data.
9. Understanding that the experiment needs to be repeated.
10. Awareness that one can never prove a hypothesis—that one can never be 100% sure; that there might be another experiment that could be done that would disprove the hypothesis; that there are possible sources of error; that there are limits to generalizing the conclusions (credit for any of these).

Student EDAT responses were randomized among scorers each year to avoid any bias that could result from faculty members scoring their own students’ responses. Scorer reliability was established by having all scorers evaluate a subset of EDATs from a prior introductory biology course at Doane where EDAT data had been collected (but was not part of the dataset in this study). A common set of six faculty members scored EDATs according to an established, internal rubric based on discussions of the initial subset of student responses described above. All scores across CURE (Brome and SEA-PHAGES), year, and section/instructor were aggregated into a single dataset where changes in EDAT score (i.e., paired response of post-test vs. pre-test score) could be compared.

**Statistical analysis of EDAT data**

Only EDAT scores for students where both pre- and posttest data were available were used in the final statistical analysis described below (i.e., students who were absent on the day of the EDAT or who dropped the course before the end of the semester did not have paired scores). Of the 346 students’ responses in the original dataset, 63 students’ scores were removed for missing either a pre- or posttest score (18.2% of all student responses were removed from the dataset). For students with both pre- and posttest scores, we computed a single response metric, change in score (Post EDAT – Pre EDAT evaluation; N = 283 students).

Prior to analysis, we normalized changes in students’ scores to account for the amount of room for improvement between assessments [(Post EDAT – Pre EDAT)/(Max Possible (10) – Pre EDAT)]. To compare change in score between the Brome CURE and the SEA-PHAGES CURE, we used a linear mixed-effects (LME) model in which section was held as a random effect (N = 18 sections; 10 Brome CURE and 8 SEA-PHAGES CURE). We held section as a random variable for two reasons: 1) variation within a section is likely to differ from variation among all students (i.e., students within a section are not truly independent replicates), and 2) we were not interested in the variation among sections (18).

Year (2013;2015), system (Brome CURE and SEA-PHAGES CURE), and the interaction between year and system were included as fixed factors to determine whether students’ scores differed with time or system and/or were affected differently in each system over time.

To determine which scoring criteria within the EDAT were most diagnostic for explaining changes in scores between systems, we performed a post-hoc discriminant power analysis comparing changes in scores for each question between sections (19). By utilizing both the change in overall EDAT score and changes in individual EDAT criteria, we can see whether the Brome CURE does in fact help students better understand particular parts of the experimental design process while also identifying parts of the process that are not addressed well by the Brome CURE.

Overall, our analysis allowed us to determine the impact of the Brome CURE on changes in EDAT score independent of year and section and to compare those changes with EDAT data from the SEA-PHAGES CURE. All statistical analyses were performed using the R statistical program (R Core Team, 2016). Approval for this study was obtained from the Doane University Institutional Review Board (approval no. F16 EX01 DC IRB HS).

**Other aspects of CURE implementation**

Although the specific CUREs differ, all sections used some parallel implementation strategies that are cited as best practices for a CURE (4, 7, 8), including (but not limited to) students working in groups on their investigation, open-ended outcomes, peer-to-peer collaboration, statistical analyses of data, peer assessment, not change and the EDAT was administered in the same manner across all years and sections.

In both CUREs, students were given the EDAT prompt on the first day of the course with minimal instruction, and the students were assessed again with a similar prompt at the end of the course (13). As recommended by Sirum and Humburg (13), the pretest EDAT scores and the EDAT scoring rubric were not shared with students during the semester. The post-test was administered in class at the end of the semester, before the final exam. Students were given a small amount of course credit for completing each test (pre and post).

The EDAT prompt invites students to describe an investigational design based on a claim about a health supplement. This prompt does not require the students to have any prior knowledge of the claim and is independent of the content in both CUREs. Both pre- and post-tests were scored by members of the Doane University biology department using the 10 scoring criteria described by Sirum and Humburg (13):

1. Recognition that an experiment can be done to test the claim.
2. Identification of what variable is manipulated.
3. Identification of what variable is measured.
4. Description of how the dependent variable is measured.
5. Understanding of the placebo effect.
6. Realization that there is one other variable that must be held constant (versus no mention).
7. Realization that there are many variables that must be held constant (versus only one or no mention).
8. Understanding that the larger the sample size or number of subjects, the better the data.
9. Understanding that the experiment needs to be repeated.
10. Awareness that one can never prove a hypothesis—that one can never be 100% sure; that there might be another experiment that could be done that would disprove the hypothesis; that there are possible sources of error; that there are limits to generalizing the conclusions (credit for any of these).

Student EDAT responses were randomized among scorers each year to avoid any bias that could result from faculty members scoring their own students’ responses. Scorer reliability was established by having all scorers evaluate a subset of EDATs from a prior introductory biology course at Doane where EDAT data had been collected (but was not part of the dataset in this study). A common set of six faculty members scored EDATs according to an established, internal rubric based on discussions of the initial subset of student responses described above. All scores across CURE (Brome and SEA-PHAGES), year, and section/instructor were aggregated into a single dataset where changes in EDAT score (i.e., paired response of post-test vs. pre-test score) could be compared.

**Statistical analysis of EDAT data**

Only EDAT scores for students where both pre- and posttest data were available were used in the final statistical analysis described below (i.e., students who were absent on the day of the EDAT or who dropped the course before the end of the semester did not have paired scores). Of the 346 students’ responses in the original dataset, 63 students’ scores were removed for missing either a pre- or posttest score (18.2% of all student responses were removed from the dataset). For students with both pre- and posttest scores, we computed a single response metric, change in score (Post EDAT – Pre EDAT evaluation; N = 283 students).

Prior to analysis, we normalized changes in students’ scores to account for the amount of room for improvement between assessments [(Post EDAT – Pre EDAT)/(Max Possible (10) – Pre EDAT)]. To compare change in score between the Brome CURE and the SEA-PHAGES CURE, we used a linear mixed-effects (LME) model in which section was held as a random effect (N = 18 sections; 10 Brome CURE and 8 SEA-PHAGES CURE). We held section as a random variable for two reasons: 1) variation within a section is likely to differ from variation among all students (i.e., students within a section are not truly independent replicates), and 2) we were not interested in the variation among sections (18).

Year (2013;2015), system (Brome CURE and SEA-PHAGES CURE), and the interaction between year and system were included as fixed factors to determine whether students’ scores differed with time or system and/or were affected differently in each system over time.

To determine which scoring criteria within the EDAT were most diagnostic for explaining changes in scores between systems, we performed a post-hoc discriminant power analysis comparing changes in scores for each question between sections (19). By utilizing both the change in overall EDAT score and changes in individual EDAT criteria, we can see whether the Brome CURE does in fact help students better understand particular parts of the experimental design process while also identifying parts of the process that are not addressed well by the Brome CURE.

Overall, our analysis allowed us to determine the impact of the Brome CURE on changes in EDAT score independent of year and section and to compare those changes with EDAT data from the SEA-PHAGES CURE. All statistical analyses were performed using the R statistical program (R Core Team, 2016). Approval for this study was obtained from the Doane University Institutional Review Board (approval no. F16 EX01 DC IRB HS).

**Other aspects of CURE implementation**

Although the specific CUREs differ, all sections used some parallel implementation strategies that are cited as best practices for a CURE (4, 7, 8), including (but not limited to) students working in groups on their investigation, open-ended outcomes, peer-to-peer collaboration, statistical
analysis, and reading of primary literature. Additionally, professors taught both CURES, and those teaching the SEA-PHAGES CURE all went through a common one-week training course held at HHMI prior to the initial implementation of SEA-PHAGES. The Brome CURE instructors collaborated in building the syllabus for the Brome CURE (see Appendix 1). In both CURES, students are initially introduced to data collection methods and basic study design through a common activity focused on measurements of human behavior where students decide on a common set of metrics. This activity was carried out during the first week of the course. Lastly, students in both CURES participated in an end-of-semester poster session as a form of summative assessment. At this session, students presented posters describing their semester-long projects to their peers (i.e., other students, faculty, and members of the local community) in the format of a professional scientific meeting.

The Brome CURE was developed by the authors at Doane University specifically for introductory-level students and has not been previously described in the literature. The Brome CURE includes a number of other facets, not formally assessed by the EDAT, which makes it an ideal CURE system; these are summarized in Table 1. Furthermore, the focal species in this CURE (invasive species Bromus inermis; non-invasive “control” species Lolium perenne) are robust to student error (e.g., students accidentally forgetting to water the plants for a few days). Additionally, first-year students could easily understand the ecological significance of invasive species and there is a robust, highly accessible body of peer-reviewed literature on both invasive species in general and Bromus inermis in particular (20–23). For students in the Brome CURE, the combination of existing primary literature and the results of their own primary literature searches helped students develop their own hypotheses and experiments. Additionally, many of the response variables found in the primary literature (4, 23) are the same as those that students measured in their own experiments, which makes for a strong connection between student research and published research. The primary literature provided to the students in the Brome CURE system served as the basis for more detailed discussions of experimental design, hypotheses, response variables, and statistical analyses that the students utilized during their experiments. In addition, students in the Brome CURE carried out statistical analyses (via the statistical program R-commander) on a variety of synthesized datasets during the course of the semester. These datasets were either collected by the students themselves or were provided by the instructor (see Appendix 1). The analyses that the students carried out on these “practice” datasets paralleled a number of the analyses that they ultimately carried on their experimental data. These analyses included analysis of variance (ANOVA) and regression and correlation analysis. Due to the nature of the data collected in the SEA-PHAGES CURE, there was less emphasis on students there carrying out statistical analysis; however, students in the SEA-PHAGES CURE were introduced to the basic ideas of ANOVA, regression, and correlation analysis in order to understand the results of the primary literature that they were exposed to in that CURE (see Appendix 2).

RESULTS

Students in the Brome CURE and SEA-PHAGES CURE sections did not differ significantly in pretest EDAT scores (ANOVA: F = 0.172, p = 0.679), with Brome students averaging 3.74 (±0.12SE) and SEA-PHAGES students averaging 3.84 (±0.14SE). Students in the Brome CURE sections showed a significantly greater increase in EDAT scores over
the course of the semester than did students in the SEA-PHAGES CURE, as measured by the normalized change in EDAT score \((t = -3.43, p = 0.003; \text{Fig. 1})\). In the Brome CURE sections, EDAT scores increased by 15.5% on average, while scores in the SEA-PHAGES sections increased by 3.6% (Fig. 1). The Brome CURE posttest average score was 5.29 (±0.011SE) and the SEA-PHAGES CURE posttest average score was 4.20 (±0.013SE). Year had no significant effect on the change in score \((t = -0.349, p = 0.732)\), nor was there a significant interaction between system and year on the change in score \((t = 1.009, p = 0.330)\).

There were particularly large gains in the proportion of Brome students including EDAT scoring criteria 4, 6, 7, 8, and 9 in their responses. For example, the number of students receiving credit for EDAT criteria 8 (“Understanding that the larger the sample size or number of subjects, the better the data”) increased from less than 20% at the beginning of the semester to nearly 60% at the end of the semester (Fig. 2A, B). The results indicate that Brome CURE students had significantly higher gains than SEA-PHAGES students in 6 of the 10 criteria (Fig. 2C).

**DISCUSSION**

The goal of this work was to introduce the Brome CURE and to assess its effectiveness in addressing the Vision and Change competency related to the application of the scientific process through the development and testing of hypotheses (Fig. 3). We showed that first-year biology
FIGURE 3. Schematic of the implementation of the Brome CURE and EDAT. Also pictured are extensions not measured by the EDAT that address other Vision and Change competencies and student-driven positive feedbacks on the Brome CURE. Vision and Change Competencies: 1) application of the scientific process through development and testing of hypotheses; 2) use of quantitative reasoning; 3) use of modeling and simulation; 4) recognition and utilization of the interdisciplinary nature of science; 5) communication with other scientific disciplines; and 6) understanding the relationship between science and society. V&C = Vision and Change; EDAT = Experimental Design Ability Test.

students participating in the Brome CURE did indeed meet this competency, as reflected in the improvement of their EDAT scores relative to the scores of students in the nationally implemented SEA-PHAGES CURE. These gains are quantitatively consistent with other studies that reported gains in EDAT score with the use of a CURE and CURE-like activities (13, 17). Furthermore, Brome CURE students had significantly higher gains relative to SEA-PHAGES students in 6 of the 10 EDAT scoring criteria. The largest differences were found in criteria focused on sample size, dependent variable identification and measurement, and experimental controls, which probably reflects differences in methodologies and scope between the two CUREs.

Differences in how students approach their data and experimental design in each CURE may account for the differential gains in EDAT scores between the two CUREs. In the Brome CURE, data are more quantitative in nature, which not only provides an opportunity for students to be introduced to basic statistical analysis (e.g., ANOVA, regression, correlation) and the development of quantitative thinking skills, but may also lead to a more robust grasp of concepts such as independent and dependent variables. Data collected in the SEA-PHAGES CURE, on the other hand, are more qualitative in nature (e.g., presence/absence of bacteriophage in soil samples).

The EDAT assessment measures a specific component of first-year biology students’ educational experience, namely experimental design and hypothesis testing. While other assessments focusing on different educational aspects would be of interest (24–26), examining additional aspects was beyond the intended scope of our study. For example, while we did not collect formal data on students’ attitudes toward science, anecdotally, we found that SEA-PHAGES CURE students felt that they were “doing science” because they were developing a number of more traditional lab skills (e.g., pipetting and streak plating), conducting more bench work, and investigating a microscopic biological phenomenon.

Both CUREs involve or allow for other activities and extensions that enhance student understanding of science in general and are not measured by the EDAT (e.g., reading primary literature and communicating scientific ideas to the public). This is particularly relevant because research has shown that exposure to primary literature in a structured manner (as done in both CUREs) can lower barriers to understanding science that students often feel when presented with primary literature for the first time (27, 28). Furthermore, the summative end-of-semester poster session provided a professional development opportunity, allowing students to see how their original work fits into a larger scientific context, and required them to organize their findings into a logical and accessible sequence. These types of assessments can in turn help students develop their ability to communicate with individuals in other disciplines (Vision and Change Competency 5) (29, 30). Furthermore, the poster session allowed students in both CURE systems to interact with each other and ask questions about their respective semester-long projects. Lastly, both systems have resulted in extensions of the research projects beyond the introductory course.

Overall, based on our analysis of the EDAT data, the Brome CURE helps students develop a meaningful understanding of some of the competencies outlined in Vision and Change, particularly experimental design and application of the scientific process through the development and testing of hypotheses. This was clearly demonstrated in our data by the significant increase in the EDAT scores of Brome CURE students. Additionally, the high level of accessibility of the Brome CURE allows it to be used in diverse experimental contexts, providing an opportunity for students to take ownership of the scientific process in a way that “cookbook” laboratory exercises do not (5). The Brome CURE
can easily be adopted into introductory biology courses to give students the opportunity to conduct self-designed experiments that address the real-world problem of invasive species. Furthermore, these introductory experiments can extend beyond a semester-long course into more nuanced projects (e.g., summer or senior research projects) to help students gain an even deeper understanding of the scientific process through the development and testing of hypotheses and the use of quantitative reasoning—key competencies outlined in Vision and Change. The Brome CURE provides biology departments with an effective, ecological parallel to the microbial SEA-PHAGES CURE. Furthermore, if it is possible to run both systems in tandem across different sections of an introductory biology course (as we did here), students will be provided with contrasting examples of the scientific process at two different biological scales.

SUPPLEMENTAL MATERIALS

Appendix 1: Brome CURE course schedule/syllabus
Appendix 2: SEA-PHAGES CURE course schedule/syllabus

ACKNOWLEDGMENTS

First, we would like to thank all the students who were part of the dataset and congratulate them on the amazing science that they did in both the Brome CURE and SEA-PHAGES CURE. We would also like to acknowledge Dr. Chad Brassil and Dr. Brian Couch for their insightful comments on earlier versions of the manuscript and Dr. Tim Frey for evaluating the research for Doane’s Institutional Research Board. Training support for SEA-PHAGES was provided by HHMI. The authors declare that there are no conflicts of interest.

REFERENCES

2. National Research Council (US) Committee on Undergraduate Research Board. Training support for SEA-PHAGES was provided by HHMI. The authors declare that there are no conflicts of interest.