

Multi-Institutional, Multidisciplinary Study of the Impact of Course-Based Research Experiences[†]

Catherine M. Mader^{1*}, Christopher W. Beck², Wendy H. Grillo³, Gail P. Hollowell³, Bettye S. Hennington⁴, Nancy L. Staub⁵, Veronique A. Delesalle⁶, Denise Lello⁷, Robert B. Merritt⁷, Gerald D. Griffin^{8,10}, Chastity Bradford⁸, Jinghe Mao⁴, Lawrence S. Blumer⁹, and Sandra L. White¹¹

¹Department of Physics, Hope College, Holland, MI 49423, ²Department of Biology, Emory University, Atlanta, GA 30322, ³Department of Biology, North Carolina Central University, Durham, NC 27707,

⁴Department of Biology, Tougaloo College, Tougaloo, MS 39174, ⁵Biology Department, Gonzaga University, Spokane, WA 99258, ⁶Department of Biology, Gettysburg College, Gettysburg, PA 17325, ⁷Department of Biological Sciences, Smith College, Northampton, MA 01063, ⁸Department of Biology, Tuskegee University,

Tuskegee Institute, AL 36088, ⁹Department of Biology, Morehouse College, Atlanta, GA 30314,

¹⁰Department of Biology, Hope College, Holland, MI 49423, ¹¹Center for Science, Math and Technology Education, North Carolina Central University, Durham, NC 27707

Numerous national reports have called for reforming laboratory courses so that all students experience the research process. In response, many course-based research experiences (CREs) have been developed and implemented. Research on the impact of these CREs suggests that student benefits can be similar to those of traditional apprentice-model research experiences. However, most assessments of CREs have been in individual courses at individual institutions or across institutions using the same CRE model. Furthermore, which structures and components of CREs result in the greatest student gains is unknown. We explored the impact of different CRE models in different contexts on student self-reported gains in understanding, skills, and professional development using the Classroom Undergraduate Research Experience (CURE) survey. Our analysis included 49 courses developed and taught at seven diverse institutions. Overall, students reported greater gains for all benefits when compared with the reported national means for the Survey of Undergraduate Research Experiences (SURE). Two aspects of these CREs were associated with greater student gains: 1) CREs that were the focus of the entire course or that more fully integrated modules within a traditional laboratory and 2) CREs that had a higher degree of student input and results that were unknown to both students and faculty.

INTRODUCTION

Numerous calls have been made to reform undergraduate science education by engaging more students directly in scientific research within the curriculum (1, 2, 3). The *Engage to Excel* report (4) highlights the need to “replace the standard laboratory courses with discovery-based research courses.” In response to these reports, course-based research experiences, or CREs, have been developed to engage undergraduates in research, especially

at institutions that cannot accommodate large numbers of undergraduates in apprentice-style research (5, 6). Engagement in undergraduate research has been widely credited as an effective mechanism for enhancing the learning experience of undergraduates (7). By bridging scientific research and education, CREs help students to better understand what scientists do, learn to think like scientists, become enthused by doing research, and reinforce their decisions to pursue science, technology, engineering, and mathematics (STEM) careers (8–12). CRE experiences not only provide the opportunity to involve more students in research, but also provide this opportunity to a greater diversity of students (13). CREs involve entire classes of students working on research projects that are of interest to the scientific community and allow the students to engage in scientific practices, such as asking questions, gathering and analyzing data, and communicating their findings (8, 14, 15).

*Corresponding author. Mailing address: Department of Physics, Hope College, Holland MI 49423. Phone: 616-395-7114. Fax: 616-395-7123. E-mail: mader@hope.edu.

Received: 8 March 2017, Accepted: 13 June 2017, Published: 1 September 2017.

[†]Supplemental materials available at <http://asmscience.org/jmbe>

Recent reviews of inquiry-based learning and course-based research in biology laboratory courses suggest that CREs can lead to significant gains by students in self-efficacy, disciplinary content knowledge, analytical, technical, and science process skills, understanding of the nature of science, persistence in science, and career clarification (7, 16–18). However, in many cases, only a single course at a single institution with unique assessment practices is considered (16). As a result, Lopatto (19) argues for a “consortium model” in which CREs are assessed across a range of institution types. These consortium models have included a large number of institutions which have all implemented the same CRE model (see Table 2 in Elgin et al. [20]). Whether the results from these studies can be generalized to all CRE models at all institution types remains an open question, and research that examines student outcomes using the same assessment instrument across a range of courses and institutions is needed (16). Furthermore, which components of CREs (8) lead to the greatest student gains has not been examined in detail (7). Preliminary research shows that course elements common to high research courses (e.g., student input into the experimental design, student research proposals, student oral presentations of results) have led to greater student self-reported gains (19).

The analysis presented here builds on this preliminary research and explores the impact of different CRE models in different contexts on student self-reported gains in understanding, skills, and professional development. The analysis includes courses developed and taught at seven institutions, public and private, universities and liberal arts colleges, with support from HHMI (Howard Hughes Medical Institute) Undergraduate Education Awards. This seven-institution team is referred to as the CRE Collaboration (see 15). The diversity of implementation of CREs across these institutions allowed us to begin to address whether student gains as a result of CREs are generalizable, and which CRE structures (e.g., module vs. full semester) and components (e.g., reading primary literature, presenting results) lead to the greatest student gains. Within the CRE Collaboration, each CRE had unique goals and contexts. Thus, each of the courses was developed and implemented using different models to fit the institutional curricular structure, faculty interests, student population, and resources. Additional details on the institutions and CRE courses offered at the institutions can be found in Staub et al. (15).

Prior to the establishment of the Collaboration, most institutions had developed program assessment plans that coincidentally included use of the Classroom Undergraduate Research Experience (CURE) survey or the closely related Research on the Integrated Science Curriculum (RISC) survey. These surveys identify self-reported gains in understanding related to the nature of science, in skills related to doing science, as well as in academic and professional development as a scientist after a CRE experience. These are often referred to as “benefits” (Appendix 2).

Although assessments that are based on student self-reports have been criticized (e.g., 18), other methods to measure these types of student outcomes across diverse courses with different student populations and different learning goals were not available when this collaborative project was initiated (21). The CRE developers on collaboration campuses who participated in the present study did not rely only on student self-reported gains to determine that students are developing skills, abilities, and understanding. The instructional staff also evaluated student work in traditional manners (tests, laboratory notebooks, presentations, and papers) in order to measure student development. However, the student self-reported results of the CURE/RISC surveys provide a common measure of the perceptions of students participating in CREs across a wide range of course models. By having a common measure, we were able to explore the impact of different CRE models on student outcomes.

METHODS

Data collection

Each of the 49 CRE offerings included in this study, that affected more than 1,350 students, used the Research Experience Benefits component of the CURE/RISC surveys as a part of their course assessment plan. Since the institutions used the CURE/RISC surveys in different ways, each one followed its own IRB approval process for use of these instruments. In addition, IRB exemption was obtained for use of the existing results from each institution for this work. Multiple sections of a course taught at a single institution in a specific semester were reported as a single course. However, the same course offered in different semesters was treated as a separate CRE offering. Students in all courses were asked to report gains in 20 potential benefits of research experiences at the end of their course-based research experience. Several courses were full-year courses, and students completed the survey at the end of the year-long course. Course averages for each benefit were collected and used as a part of the 49-course sample.

To identify the different characteristics of each CRE model, faculty involved in the development and teaching of these CRE models completed several surveys. In one survey, faculty identified the context of their course, including the scientific discipline and the student audience. Each course was also identified with respect to the extent of the research experience as follows:

- a full-course focused on one research experience
- a sequence of short modular research experiences within a single laboratory course
- an interlude in which a research module is embedded within a more traditional laboratory course or
- a research module interwoven within a more traditional laboratory course.

For each of the 49 CREs, faculty members involved in teaching the course completed a course element survey based on elements included in the CURE surveys (Appendix 1). In all cases, the survey was completed in spring 2016, after the course was taught at least one time. For each possible course element, faculty members were asked to indicate “how much relative emphasis (time on task) is placed on each element in this course” by selecting from the following six options:

- Not applicable
- No emphasis
- Little emphasis
- Some emphasis
- Much emphasis
- Extensive emphasis

If multiple instructors completed the survey for a single course, a course average was determined. If a course was offered in multiple semesters, instructional staff had the option of completing additional surveys to take into account changes in emphasis over time.

Statistical analysis

For each of the 20 student self-reported benefits, we compared the average benefit of the courses in our collaboration with the national average for the Research Experience Benefits section of the Survey of Undergraduate Research Experiences (SURE), which undergraduates typically take part in after completing a traditional apprentice-model research experience. We did not compare our results with the national average for the CURE survey, as the course context in which students complete the CURE survey for the national participant pool is unclear. We calculated a standardized difference (Cohen's *d*), using the national SURE as the control group (22, 23).

To determine the effect of course structure on student self-reported benefits, we compared the average benefits between full-semester courses and module-based courses using two-level mixed effects general linear models, with institution and course as the two levels. Course structure was considered a fixed effect. Institution was considered a random effect to control for the non-independence of courses within a particular institution. In addition, we calculated standardized differences (Cohen's *d*) using a pooled standard deviation. We used the same approach to compare the average benefits between full-semester courses and the three different types of modules (sequence, interwoven, and interlude) to determine whether the way in which CRE modules are implemented affects the students' perceptions of the benefits of CREs. When the main effect of course structure was significant, we carried out post-hoc pairwise comparisons using least significant difference tests on estimated marginal means. Because we considered 20 benefits, the likelihood of finding significant effects of course structure on a particular benefit by chance alone is high. As a result, we adjusted significance levels downward using a sequential Bonferroni approach.

In addition to the effect of course structure on student perceptions of benefits in CREs, we wanted to determine whether particular course elements combine to result in differences in benefits. In our sample, the course elements were highly intercorrelated, which prevented us from using them as independent predictors in our linear models. In addition, it seems unlikely that a single course element will have a significant effect on benefits. As a result, considering which course elements covary in courses and whether they affect benefits is more important. Therefore, we used cluster analysis to group courses that were more similar in terms of the levels of emphasis on specific course elements. A similar approach was used in a recent analysis of instructional practices in chemistry laboratory courses (24). Hierarchical clustering with bootstrap resampling ($n = 1,000$

TABLE 1.
CRE models that report CURE/RISC/SURE self-assessment results in the literature.

Course	CRE Type	Reference
Genomics Education Partnership	full course	(27)
SEA-PHAGES	full course	(26)
UCLA Undergraduate Research Consortium in Functional Genomics (URCFG)	full course	(28, 29)
Phage Discovery	full course	(30)
Zebrafish Introductory Biology	full course	(31)
Microbial Genome Annotation	full course(s)	(32)
Insect Evolutionary Genetics	full course	(33)
Dynamic Genome	full course	(34)
Principles of Genetics – Maize Module	interlude	(35)
Collaborative CREs visualizing biological processes	full course(s)	(36)

CRE = course-based research experience; CURE = classroom undergraduate research experience; RISC = research on the integrated science curriculum; SURE = survey of undergraduate research experiences.

bootstraps) was carried out using the pvcust package in R (25). For the cluster analysis, we used 21 of the 25 course elements in the faculty course element survey of CURE. We excluded “Listen to lectures,” “Read a textbook,” “Work on problem sets,” and “Take tests in class” as these course elements are not common in laboratory courses.

The resulting cluster diagram (Fig. 1) shows one course as an outgroup and two larger, moderately supported clusters (“approximately unbiased” bootstrap value of 75/100). Further clustering was not supported by bootstrap values or the small number of courses within a cluster. Therefore,

we excluded the outgroup course from subsequent analyses and assigned the remaining courses to one of the two larger clusters. To determine which course elements differed between the clusters, we used two-level mixed effects general linear models with cluster as a fixed effect and institution as a random effect. We also calculated standardized differences (Cohen’s *d*) using a pooled standard deviation. The same approach was used to test for an effect of cluster on each of the 20 student self-reported benefits. Again, for both analyses, we adjusted significance levels downward using a sequential Bonferroni approach.

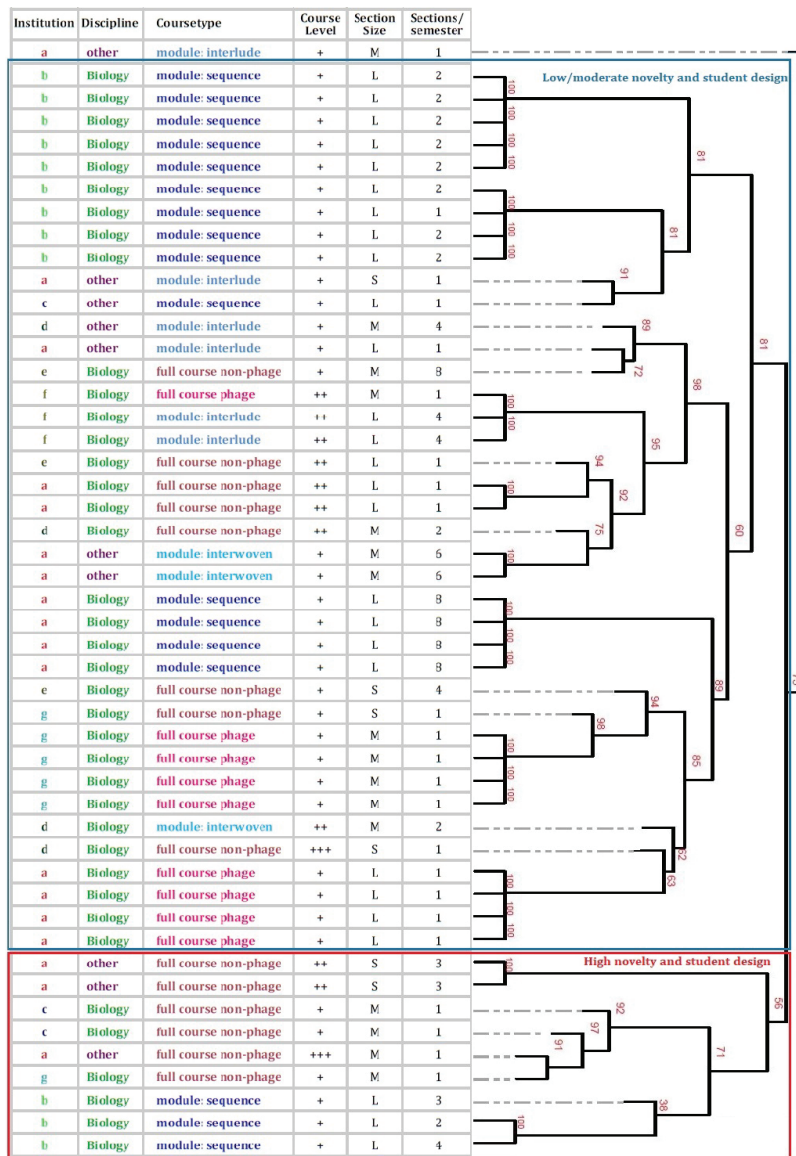


FIGURE 1. Cluster dendrogram based on faculty course elements. Values at nodes represent bootstrap values from 1,000 bootstrap samples. Institution, discipline, and course type (see text for explanation of course types) are indicated for each course. Course levels are introductory (+), intermediate (++), or advanced (+++). CREs with enrollments less than 10 (S), greater than 20 (L), or in between (M) are indicated. Courses were divided into two clusters as indicated for subsequent analysis. Based on differences in the course elements between the clusters (see Fig. 5), we defined courses as either “Low/moderate novelty and student design” or “High novelty and student design.” Smaller clusters represented as polytomies with high bootstrap support (100/100) are the same courses taught across multiple semesters. CRE = course-based research experience.

RESULTS

Comparison of CREs with national SURE results

As a benchmark to other programs, we compared the average gains in the CURE survey benefits for the CREs in our collaboration to the national averages for students completing the SURE after finishing apprentice-style research experiences (Fig. 2, Appendix 2). Overall, students in collaboration CREs reported similar or greater gains for 16 benefits when compared with the national means for the SURE. In comparison with the national SURE data, the standardized differences for our collaboration CREs ranged from -0.08 (“Learning to work independently”) to 0.30 (“Ability to analyze data and other information”). Overall, the largest standardized differences were for items grouped in the categories of “Understanding” and “Skills and Abilities” (Appendix 2). Furthermore, our students noted the highest gains in “Learning laboratory techniques,” “Ability to analyze data and other information,” and “Understanding of how scientists work on real problems.” In contrast, for the national SURE data, the greatest gains were in “Understanding the research process,” “Tolerance for obstacles faced in the research process,” and “Understanding of how scientists work on real problems.” Students in our CREs

reported the lowest gains in “Clarification of a career path,” “Learning ethical conduct in your field,” and “Skill in how to give an effective oral presentation”—the only benefits with gains between 3.0 and 3.5 on the 5-point Likert scale. The benefits with the lowest gains for the national SURE data were similar, but include “Skill in science writing” rather than “Skill in how to give an effective oral presentation.”

Comparison of different course types

Students who participated in CREs that were full courses reported significantly greater gains than students who participated in CREs that were modules (mixed effects general linear models [GLMs]; Fig. 3). The standardized differences between full-course and module-based CREs ranged from 0.567 (“Learning ethical conduct”) to 1.809 (“Readiness for more demanding research”) (Appendix 3). Although students who participated in modular CREs reported lower gains than students in full-course CREs, the benefits of modular CREs were typically comparable to the gains reported in the national SURE average. The three highest gains in benefits in the full courses were “Learning laboratory techniques,” “Ability to analyze data and other information,” and “Ability to read and understand primary literature.” Students in the module-based CRE reported the highest gains in the same

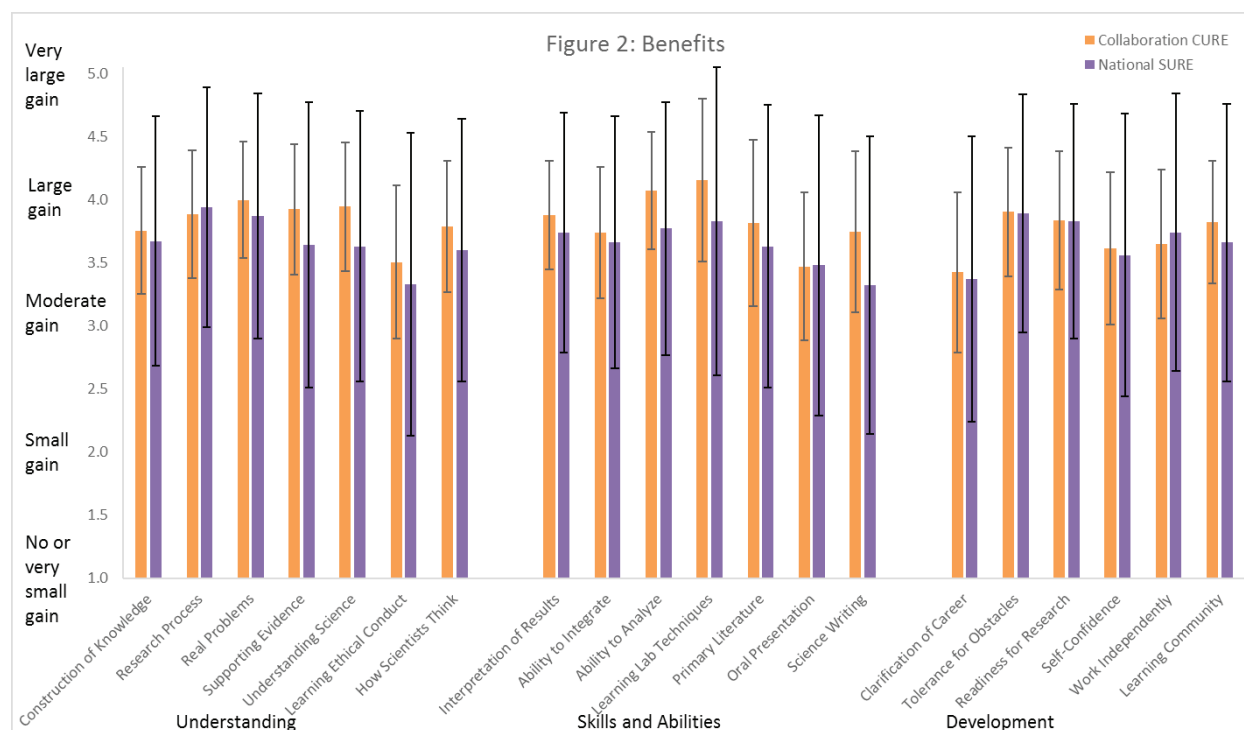


FIGURE 2. Student self-reported benefits from CURE survey and SURE results. The Collaboration CURE mean and standard deviations represent the average of 49 different course means from collaboration CREs with standard deviations indicated. The National SURE mean and standard deviations are for summer 2014 averages for $\leq 3,041$ student responses. Overall, students in our collaboration CREs reported greater or the same gains for 16 benefits when compared with the national means for the SURE. Data are presented in Appendix 2. CRE = course-based research experience; CURE = classroom undergraduate research experience; SURE = survey of undergraduate research experiences.

benefits with the exception of “Understanding of how scientists work on real problems” replacing “Ability to read and understand primary literature.” The lowest gains in benefits for both full-course and modular CREs were “Skill in how to give an effective oral presentation,” “Learning ethical conduct in your field,” and “Clarification of a career path.” These are consistent with the lowest gains reported for the national SURE results.

In general, students who participated in courses in which modules were interwoven with more traditional laboratory exercises reported similar benefits to students in full-course CREs. This can be seen in Figure 4 (see also Appendix 4), where the means for interwoven modules are typically between the full-course and sequence-of-module formats and, as indicated by the shared letters above the interwoven module results, statistically indistinguishable from the full-course means and sequence-of-module means. However, the sequence-of-module means are statistically different from the full-course means. In contrast, the benefits of interlude modules were significantly lower than those for most other CRE formats. Whether particular approaches to implementing modules differed in their benefits depended on the specific benefit. However, the number of courses that used interlude ($N = 6$) and interwoven ($N = 3$) modules was low, so differences in the benefits of implementing CRE modules in a particular way should be interpreted with caution. Future studies could explore in

more detail how implementing CRE modules in different ways results in different student outcomes.

Comparison based on CRE elements

We also examined whether CREs that differed in particular course elements were associated with different student gains. Based on a hierarchical cluster analysis of course elements, the courses in our collaboration grouped into two major clusters with moderate support (Fig. 1). Both clusters included full-course and modular CREs, as well as courses from biology and other STEM disciplines. In addition, different course sizes, experience levels of the student audience, and experience of the instructors are found within each cluster. However, cluster 2 only contained full-course CREs that did not implement the SEA-PHAGE (26) model as well as CREs with modules that were implemented in a sequence.

Courses in the two clusters were significantly different in a range of course elements (Fig. 5, Appendix 5). Courses in cluster 1 tended to be more scripted and more often included projects structured by the instructor than courses in cluster 2. In contrast, CREs in cluster 2 involved greater emphasis than the cluster 1 courses on the following: student input into the design process or topic of the CRE and students that were responsible for part of the project, wrote research proposals, collected data, maintained lab notebooks, presented posters, and critiqued the work of

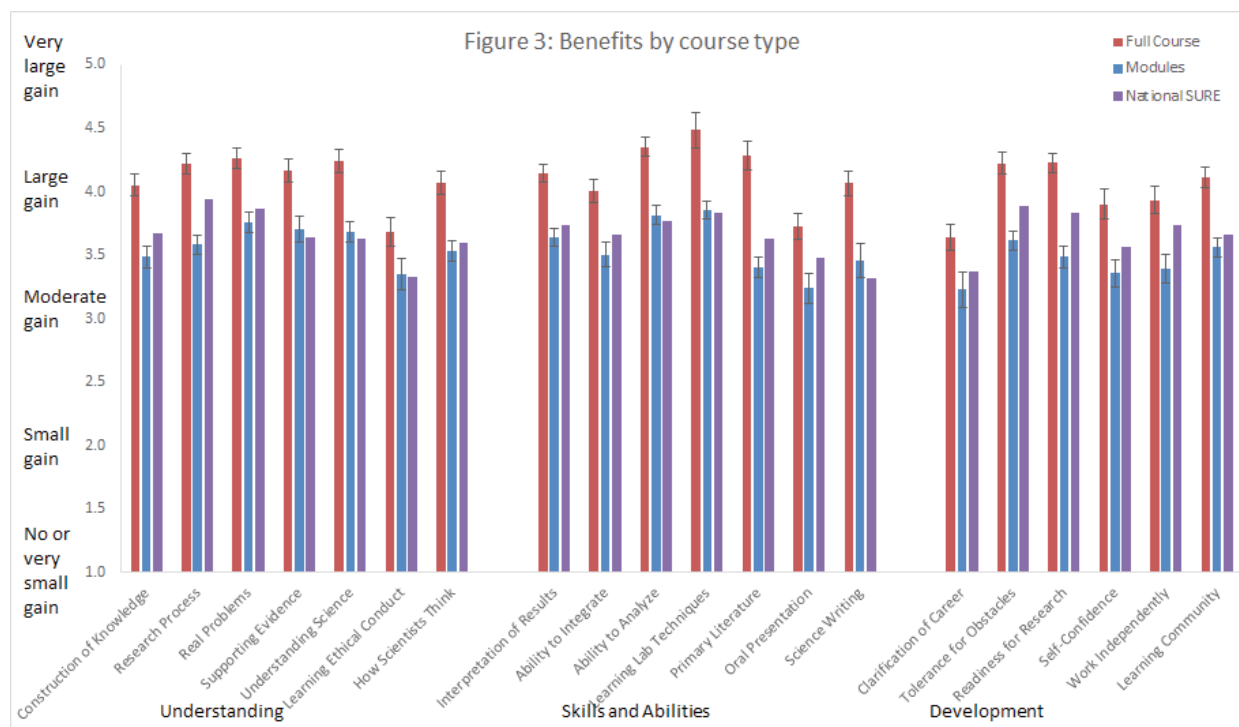


FIGURE 3. Student self-reported benefits based on whether the CRE was a full-course experience or a module within a course. Means and standard errors are shown for full courses and modules (sequence, interlude, and interwoven). Students who participated in CREs that were full courses reported significantly greater gains than students who participated in CREs that were modules (Mixed effects GLMs, all comparisons significant after controlling for experimentwise-error rate with sequential Bonferroni). Numerical data are presented in Appendix 3. CRE = course-based research experience; GLM = general linear model.

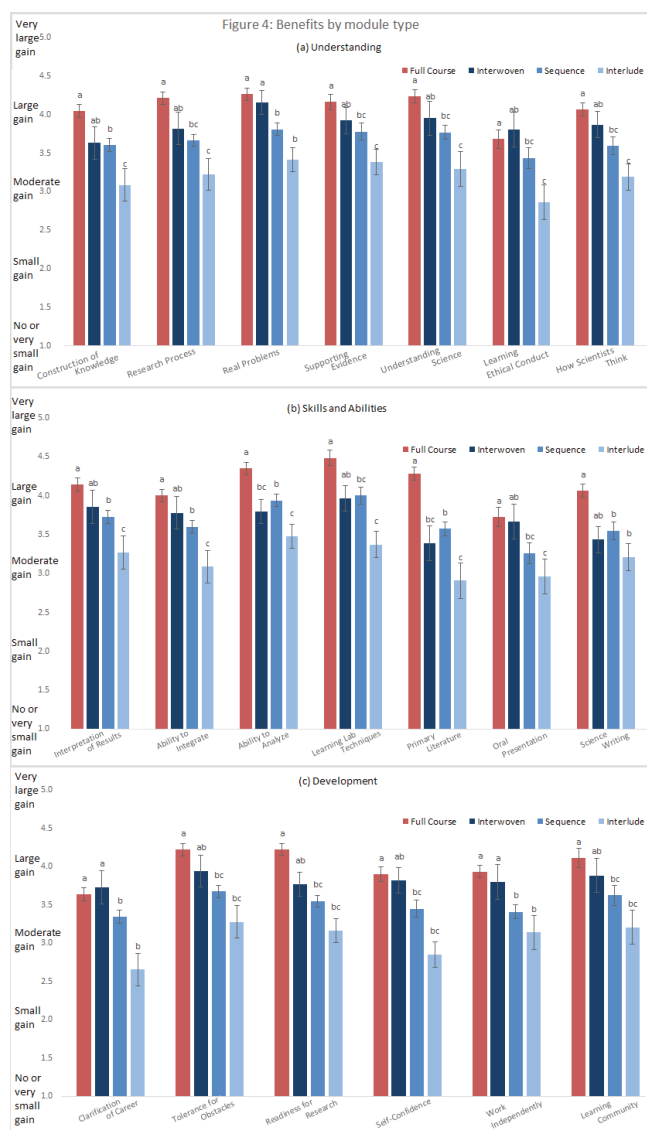


FIGURE 4. Student self-reported benefits based on module type. Means and standard errors are shown for full courses and the three module types (sequence, interlude, and interwoven). Letters above bars indicate the highest mean in each item (a) to the lowest. Shared letters indicate items are not statistically significantly different. In general, students who participated in courses in which modules were interwoven with more traditional laboratory exercises reported similar benefits to students in full-course CREs. Numerical data are presented in Appendix 4. CRE = course-based research experience.

others. In addition, CREs in cluster 2 were more often projects in which no one knew the outcome, as compared with CREs in cluster 1. The most pronounced differences between the clusters in course elements based on standardized differences were in students writing research proposals and students engaging in research projects of their own design. Based on the differences in the course elements in the two clusters, we defined cluster 1 as “Low/moderate novelty and student design” courses and cluster 2 as “High novelty and student design” courses.

Students in the “High novelty and student design” CREs reported significantly greater gains for all benefits with the exception of “Skill in how to give an effective oral presentation” (mixed effects GLMs, Fig. 6). For benefits that differed significantly between “High novelty and student design” CREs and “Low/moderate novelty and student design” CREs, standardized differences ranged from 0.568 (“Clarification of a career path”) to 1.17 (“Skill in interpretation of results”) (Appendix 6). Although students in “Low/moderate novelty and student design” CREs reported significantly lower gains in benefits than students in “High novelty and student design” CREs, the gains in benefits for “Low/moderate novelty and student design” CREs were greater than or equivalent to those for the national SURE data (Fig. 6). Interestingly, the greatest reported gains in benefits as well as the lowest reported gains in benefits were the same for both CRE types. The highest gains in benefits included “Learning laboratory techniques,” “Ability to analyze data and other information,” and “Understanding of how scientists work on real problems,” while the lowest gains in benefits were “Skill in how to make an effective oral presentation,” “Clarification of a career path,” and “Learning ethical conduct in your field.”

DISCUSSION

For our CRE collaboration, courses that completely focused on the research experience, rather than compartmentalizing the research experience into a short interval within a traditional laboratory course, showed gains that significantly exceeded the National SURE gains. Courses in which modules were used, but in ways that engaged students in research throughout the bulk of the course, by either interweaving the research activities within the traditional course or by doing a sequence of research experiences, had very similar gains to full-course experiences and, for many benefits, were not statistically different from the gains reported for full-course experiences. However, courses in which the research experience was compartmentalized into a small component (or interlude) of the course resulted in lower gains than those reported by students in the sequential or interwoven CREs in which the research experience is spread throughout the semester. Therefore, although all of the CREs in our collaboration led to student-reported gains in a range of benefits, CREs that lasted an entire semester or that more fully integrated modules led to greater student gains.

Using common pedagogical elements to group courses, collaboration CRE models fell into two clusters (Fig. 1) representing “Low/moderate novelty and student design” and “High novelty and student design.” These course components correspond to the two perspectives on authentic research experiences (Process of Science and Novel Questions) reported by Spell et al. (12). The clusters also indicate differences in the course elements that reflect efforts to engage students in discovery and the use of relevant science research practices (e.g., asking questions,

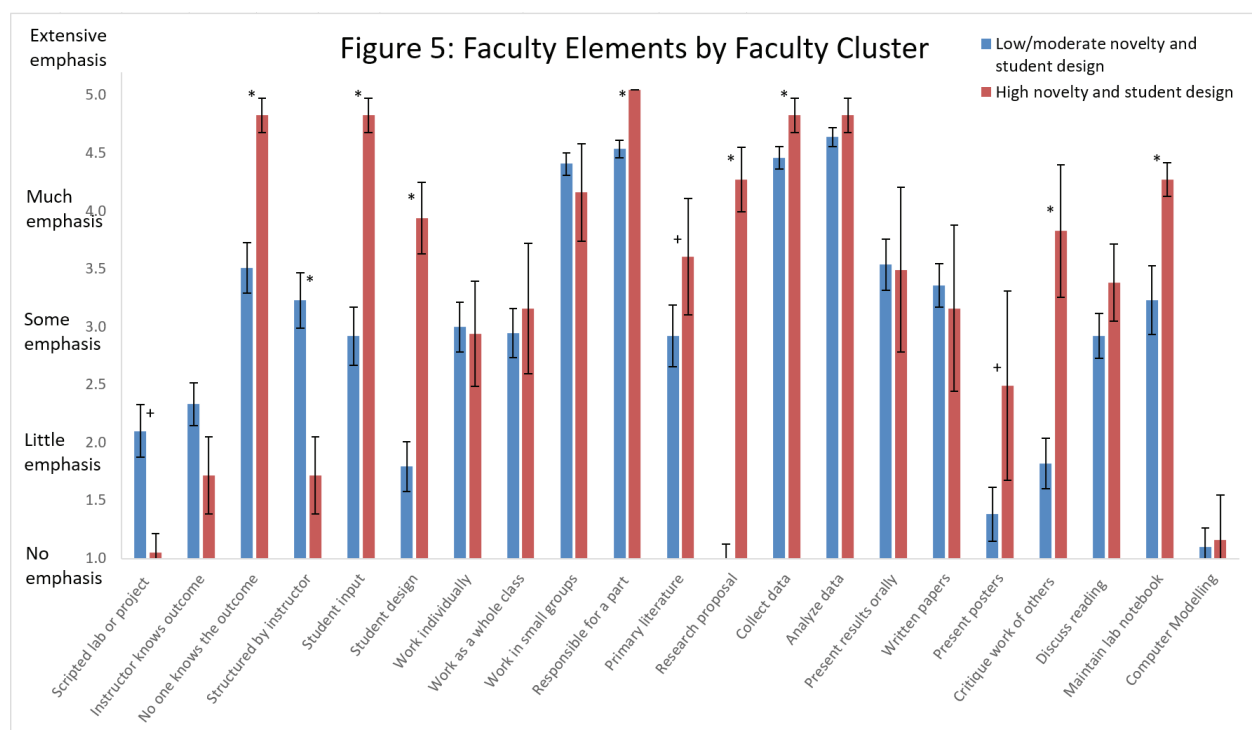


FIGURE 5. Faculty-reported emphasis of course elements based on cluster elements. Clusters are defined based on the dendrogram in Figure 1. Means and standard errors are reported for each cluster. Asterisks indicate items for which the two clusters' means are statistically significantly different after sequential Bonferroni adjustment of significance values. Crosses indicate items for which the two clusters' means are statistically significantly different at $\alpha = 0.05$, but not significantly different after sequential Bonferroni correction. Based on the differences in the course elements between the clusters, we defined courses as being either "Low/moderate novelty and student design" or "High novelty and student design." Numerical data are presented in Appendix 5.

proposing hypotheses, designing experiments, analyzing data) suggested by Auchincloss et al. (8) as components of CREs. However, these clusters do not seem to reflect different degrees of student collaboration, which was also suggested to be a significant component of CREs. Course size did not seem to be a common factor in the clusters, nor did the experience level of the student audience. In addition, the level of experience of the instructor (was this their first CRE or not) did not seem to be a common factor in the clustering.

"High novelty and student design" courses had greater student self-reported gains in the benefits than courses classified as "Low/moderate novelty and student design." While both types of CREs showed gains, courses with greater student input and unknown results ("High novelty and student design"), whether they were full-semester courses or modules, showed gains that exceed national SURE gains. For courses in which the emphasized course elements were consistent with "Low/moderate novelty and student design," the gains were consistent with those of the apprentice-model research experiences, though for two benefits ("Understanding the Research Process" and "Learning to Work Independently"), the students in these CREs reported slightly lower gains.

Student self-reported gains for most collaboration CREs exceeded the national average for apprentice-model

research experiences (as measured by SURE). Ten other CRE models have reported CURE/RISC benefit results in the literature (Table 1). In comparison with these CRE models, the gains reported by students in "High novelty and student design" CREs were consistent with or greater than most of the reported gains. The most notable exception is the Undergraduate Research Consortium in Functional Genomics program, for which gains exceed those of the collaboration in all cases. The gains for the Genomics Education Partnership, SEA-PHAGES, Dynamic Genome, and Zebrafish Introductory Biology programs are consistent with gains for most benefits, exceeding gains in a few benefits, but not in others for "High novelty and student design" CREs. All of these courses are full-semester CREs. The gains for the Collaborative CREs focusing on Visualizing Biological Processes had a range of gains across three courses which were, in most cases, consistent with the range of gains seen in both clusters in the CRE collaboration analysis, though in the case of "Clarification of a career path" and "Self-confidence," both of the CRE collaboration clusters had significantly higher gains. The gains for the remaining models included in Table 1 are lower than the gains reported by students in the "High novelty and student design" CREs for the collaboration, but they do exceed the gains reported by students in our CRE collaboration who participated in CREs that were interlude modules.



FIGURE 6. Student self-reported benefits for “Low/moderate novelty and student design” or “High novelty and student design” courses and the National SURE results. The National SURE mean and standard deviations are for summer 2014 averages for $\leq 3,041$ student responses. Means and standard errors for each item are shown. For all benefits except Oral Presentation, “High novelty and student design” courses in cluster 2 resulted in significantly higher benefits than “Low/moderate novelty and student design” courses after controlling for experimentwise-error rate with sequential Bonferroni. Numerical data are presented in Appendix 6. SURE = survey of undergraduate research experiences.

CONCLUSION

Our collaboration is the first to examine the effects of CREs over a diverse range of institutions, disciplines, course levels, and course types using a common assessment of student benefits. Overall, the CREs in our collaboration resulted in benefits that are comparable with the national average for the SURE, which suggests that CREs are a viable option for expanding access to research experiences. Unlike previous studies, we were able to examine how course structure and course elements affect learning gains (7). In general, we found that full-semester CREs, modular CREs that are fully integrated throughout the semester, and CREs that allow for student input and unknown results led to significantly greater student self-reported gains in a wide range of benefits of CREs than CREs that were not structured in these ways.

Future scale-up studies of CREs should consider student outcomes other than student self-reports (7, 18) to provide a deeper understanding of how CREs lead to different student outcomes. However, such studies are currently limited by the availability of broadly applicable instruments (18). Furthermore, future studies of the course components in CREs that lead to greatest gains could use the Laboratory Course Assessment Survey (LCAS) (37), as it captures all of the design features laid out in Auchincloss et al. (8),

some of which are missing from the CURE course element survey. Clearly, advances in our understanding of the best approaches to CREs in different contexts and how CREs lead to student gains in different contexts will require close collaboration among faculty who design and implement CREs and discipline-based education researchers who develop assessment instruments across STEM disciplines and institution types.

SUPPLEMENTAL MATERIALS

- Appendix 1: Survey of collaboration CRE developers/instructors on course elements
- Appendix 2: Student self-reported benefits from CURE and SURE survey results
- Appendix 3: Student self-reported benefits based on whether the CRE was a full-course experience or a module within a course
- Appendix 4: Student self-reported benefits based on module type
- Appendix 5: Faculty reported emphasis of course elements based on cluster elements
- Appendix 6: Student self-reported benefits for courses that are “Low/moderate novelty and student design” or “High novelty and student design”

ACKNOWLEDGMENTS

Financial support for the curricular changes and the collaboration programs was provided by grants from the Howard Hughes Medical Institute through the Undergraduate Science Education Program. We thank David Lopatto and Leslie Jaworski at Grinnell College for administering the CURE survey for many of the courses. We thank Barb Throop at Hope College for compiling all of the CURE data and carrying out preliminary analyses. We thank many faculty and staff who participated in the design and implementation of the CREs included in this study and the students who took these courses. The authors declare that there are no conflicts of interest.

REFERENCES

- American Association for the Advancement of Science. 2011. Vision and change in undergraduate biology education. A Call to Action: a summary of recommendations made at a national conference organized by the American Association for the Advancement of Science, July 15–17, 2009. Washington, DC.
- National Research Council. 2003. Bio2010: Transforming undergraduate education for future research biologists. The National Academies Press, Washington, DC.
- Woodin T, Carter VC, Fletcher L. 2010. Vision and change in biology undergraduate education, a call for action—initial responses. *CBE Life Sci Educ* 9:71–73.
- President's Council of Advisors on Science and Technology (PCAST). 2012. Engage to excel: producing one million additional college graduates with degrees in science, technology, engineering, and mathematics. Executive Office of the President, Washington, DC.
- Desai KV, Gatson SN, Stiles TW, Stewart RH, Laine GA, Quick CM. 2008. Integrating research and education at research-extensive universities with research-intensive communities. *Adv Physiol Educ* 32:136–141.
- Wood WB. 2003. Inquiry-based undergraduate teaching in the life sciences at large research universities: a perspective on the Boyer commission report. *Cell Biol Educ* 2:112–116.
- Corwin LA, Graham MJ, Dolan EL. 2015. Modeling course-based undergraduate research experiences: an agenda for future research and evaluation. *CBE Life Sci Educ* 14:es1.
- Auchincloss LC, Laursen SL, Branchaw JL, Eagan K, Graham M, Hanauer DI, Lawrie G, McLinn CM, Pelaez N, Rowland S, Towns M, Trautmann NM, Varma-Nelson P, Weston TJ, Dolan EL. 2014. Assessment of course-based undergraduate research experiences: a meeting report. *CBE Life Sci Educ* 13:29–40.
- Kardash CM. 2000. Evaluation of undergraduate research experience: perceptions of undergraduate interns and their faculty mentors. *J Educ Psych* 92:191–201.
- Laursen SL, Hunter A-B, Seymour E, Thiry H, Melton G. 2010. Undergraduate research in the sciences: engaging students in real science. Jossey-Bass, San Francisco, CA.
- Lopatto D. 2010. Science in solution: the impact of undergraduate research on student learning. Council On Undergraduate Research, Washington, DC.
- Spell RM, Guinan JA, Miller KR, Beck CW. 2014. Redefining authentic research experiences in introductory biology laboratories and barriers to their implementation. *CBE Life Sci Educ* 13:102–110.
- Bangera G, Brownell SE. 2014. Course-based undergraduate research experiences can make scientific research more inclusive. *CBE Life Sci Educ* 13:602–606.
- Hanauer DI, Dolan EL. 2014. The project ownership survey: measuring differences in scientific inquiry experiences. *CBE Life Sci Educ* 13:149–158.
- Staub NL, Blumer LS, Beck CW, Delesalle VA, Griffin GD, Merritt RB, Hennington BS, Grillo WH, Hollowell GP, White SL, Mader CM. 2016. Course-based science research promotes learning in diverse students at diverse institutions. *CUR Quarterly* 37:36–46.
- Beck CW, Butler A, Burke da Silva K. 2014. Promoting inquiry-based teaching in laboratory courses: are we meeting the grade? *CBE Life Sci Educ* 13:444–452.
- Brownell SE, Hekmat-Scafe DS, Singla V, Seawell PC, Imam JF, Eddy SL, Stearns T, Cyert MS. 2015. A high-enrollment course-based undergraduate research experience improves student conceptions of scientific thinking and ability to interpret data. *CBE Life Sci Educ* 14:ar21.
- Linn MC, Palmer E, Baranger A, Gerard E, Stone E. 2015. Undergraduate research experiences: impacts and opportunities. *Science* 347:1261757–1261761.
- Lopatto D. 2015. Consortium as experiment, p 101–122. *In* National Academies of Sciences Engineering and Medicine (ed.), Integrating discovery-based research into the undergraduate curriculum. The National Academies Press, Washington, DC.
- Elgin SC, Bangera G, Decatur SM, Dolan EL, Guertin L, Newstetter WC, San Juan EF, Smith MA, Weaver GC, Wessler SR, Brenner KA. 2016. Insights from a convocation: integrating discovery-based research into the undergraduate curriculum. *CBE Life Sci Educ* 15:fe2.
- Shortlidge EE, Brownell SE. 2016. How to assess your CURE: a practical guide for instructors of course-based undergraduate research experiences. *J Microbiol Biol Educ* 17:399–408.
- Cohen J. 1988. Statistical power analysis for the behavioral sciences, 2nd ed. Lawrence Erlbaum Associates, Hillsdale, NJ.
- Maher JM, Markey JC, Ebert-May D. 2013. The other half of the story: effect size analysis in quantitative research. *CBE Life Sci Educ* 12:345–351.
- Velasco JB, Knedeisen A, Xue D, Vickery TL, Abebe M, Stains M. 2016. Characterizing instructional practices in the laboratory: the laboratory observation protocol for undergraduate STEM. *J Chem Educ* 93:1191–1203.
- Suzuki R, Shimodaira H. 2013. Package 'pvclust'. Version 1.2-2.
- Jordan TC, Burnett SH, Carson S, Caruso SM, Clase K, DeJong RJ, Dennehy JJ, Denver DR, Dunbar D, Elgin SC, Findley AM. 2014. A broadly implementable research course in phage discovery and genomics for first-year undergraduate students. *mBio* 5:e01051-13.

27. Lopatto D, Alvarez C, Barnard D, Chandrasekaran C, Chung HM, Du C, Eckdahl T, Goodman AL, Hauser C, Jones CJ, Kopp OR. 2008. Genomics education partnership. *Science* 322:684–685.
28. Call GB, Olson JM, Chen J, Villarasa N, Ngo KT, Yabroff AM, Cokus S, Pellegrini M, Bibikova E, Bui C, Cespedes A. 2007. Genomewide clonal analysis of lethal mutations in the *Drosophila melanogaster* eye: comparison of the x chromosome and autosomes. *Genetics* 177:689–697.
29. Clark IE, Romero-Calderón R, Olson JM, Jaworski L, Lopatto D, Banerjee U. 2009. “Deconstructing” scientific research: a practical and scalable pedagogical tool to provide evidence-based science instruction. *PLoS Biol* 7:e1000264.
30. Staub NL, Poxleitner M, Braley A, Smith-Flores H, Pribbenow CM, Jaworski L, Lopatto D, Anders KR. 2016. Scaling up: adapting a phage-hunting course to increase participation of first-year students in research. *CBE Life Sci Educ* 15:ar13.
31. Sarmah S, Chism III GW, Vaughan MA, Muralidharan P, Marrs JA, Marrs KA. 2016. Using zebrafish to implement a course-based undergraduate research experience to study teratogenesis in two biology laboratory courses. *Zebrafish* 13:293–304.
32. Reed KE, Richardson JM. 2013. Using microbial genome annotation as a foundation for collaborative student research. *Biochem Mol Biol Educ* 41:34–43.
33. Miller CW, Hamel J, Holmes KD, Helmey-Hartman WL, Lopatto D. 2013. Extending your research team: learning benefits when a laboratory partners with a classroom. *BioScience* 63:754–762.
34. Burnette JM, Wessler SR. 2013. Transposing from the laboratory to the classroom to generate authentic research experiences for undergraduates. *Genetics* 193:367–375.
35. Makarevitch I, Frechette C, Wiatros N. 2015. Authentic research experience and “big data” analysis in the classroom: maize response to abiotic stress. *CBE Life Sci Educ* 14:ar27.
36. Kowalski JR, Hoops GC, Johnson RJ. 2016. Implementation of a collaborative series of classroom-based undergraduate research experiences spanning chemical biology, biochemistry, and neurobiology. *CBE Life Sci Educ* 15:ar55.
37. Corwin LA, Runyon C, Robinson A, Dolan EL. 2015. The laboratory course assessment survey: a tool to measure three dimensions of research-course design. *CBE Life Sci Educ* 14:ar37.