University of Kentucky
UKnowledge

Biology Faculty Publications | Biology

12-1-2018

A Low-Intensity, Hybrid Design Between a "Traditional" and a "Course-Based" Research Experience Yields Positive Outcomes for Science Undergraduate Freshmen and Shows Potential for Large-Scale Application

Thushani Rodrigo-Peiris
University of Kentucky, thushrodpeiris@uky.edu

Lin Xiang
University of Kentucky, lin.xiang@uky.edu

Vincent M. Cassone
University of Kentucky, vincent.cassone@uky.edu

Follow this and additional works at: https://uknowledge.uky.edu/biology_facpub

Part of the Educational Assessment, Evaluation, and Research Commons, Higher Education Commons, and the Life Sciences Commons

Right click to open a feedback form in a new tab to let us know how this document benefits you.

Repository Citation

Rodrigo-Peiris, Thushani; Xiang, Lin; and Cassone, Vincent M., "A Low-Intensity, Hybrid Design Between a "Traditional" and a "Course-Based" Research Experience Yields Positive Outcomes for Science Undergraduate Freshmen and Shows Potential for Large-Scale Application" (2018). *Biology Faculty Publications*. 163.
https://uknowledge.uky.edu/biology_facpub/163

This Article is brought to you for free and open access by the Biology at UKnowledge. It has been accepted for inclusion in Biology Faculty Publications by an authorized administrator of UKnowledge. For more information, please contact UKnowledge@lsv.uky.edu.

A Low-Intensity, Hybrid Design Between a "Traditional" and a "Course-Based" Research Experience Yields Positive Outcomes for Science Undergraduate Freshmen and Shows Potential for Large-Scale Application

Digital Object Identifier (DOI)
https://doi.org/10.1187/cbe.17-11-0248

Notes/Citation Information
Published in *CBE—Life Sciences Education*, v. 17, no. 4, ar53, p. 1-18.

© 2018 T. Rodrigo-Peiris et al. CBE—Life Sciences Education © 2018 The American Society for Cell Biology.

This article is distributed by The American Society for Cell Biology under license from the author(s). It is available to the public under an Attribution–Noncommercial–Share Alike 3.0 Unported Creative Commons License (http://creativecommons.org/licenses/by-nc-sa/3.0).

This article is available at UKnowledge: https://uknowledge.uky.edu/biology_facpub/163

A Low-Intensity, Hybrid Design between a “Traditional” and a “Course-Based” Research Experience Yields Positive Outcomes for Science Undergraduate Freshmen and Shows Potential for Large-Scale Application

Thushani Rodrigo-Peiris,[†] **Lin Xiang,**[‡] **and Vincent M. Cassone**[†*]
[†]Department of Biology and [‡]Department of STEM Education, University of Kentucky,
Lexington, KY 40506

ABSTRACT

Based on positive student outcomes, providing research experiences from early undergraduate years is recommended for science, technology, engineering, and mathematics (STEM) majors. To this end, we designed a novel research experience called the “STEMCats Research Experience” (SRE) for a cohort of 119 second-semester freshmen with diverse college preparatory levels, demographics, and academic majors. The SRE targeted student outcomes of enhancing retention in STEM majors, STEM competency development, and STEM academic performance. It was designed as a hybrid of features from apprenticeship-based traditional undergraduate research experience and course-based undergraduate research experience designs, considering five factors: 1) an authentic research experience, 2) a supportive environment, 3) current and future needs for scale, 4) student characteristics and circumstances, and 5) availability and sustainability of institutional resources. Emerging concepts for facilitating and assessing student success and STEM curriculum effectiveness were integrated into the SRE design and outcomes evaluation. Here, we report the efficient and broadly applicable SRE design and, based on the analysis of institutional data and student perceptions, promising student outcomes from its first iteration. Potential improvements for the SRE design and future research directions are discussed.

INTRODUCTION

Despite the national need to produce 1 million additional science, technology, engineering, and mathematics (STEM) undergraduate degree holders within the next decade, less than 40% of freshmen who declare STEM majors graduate with a STEM degree, and the first 2 years of college are the most critical in terms of STEM retention (Brainard and Carlin, 1998; President’s Council of Advisors on Science and Technology [PCAST], 2012; Dagley *et al.*, 2016). Also, the time taken to accomplish a STEM degree is of concern for STEM fields, in which only 35.1% of students graduate within 6 years compared with 48.7% in non-STEM fields (U.S. Department of Education 2009 data as per PCAST [2012]), thereby retarding the rate of production of the much-needed STEM-educated workforce. The uninspiring and unwelcoming nature of the introductory STEM courses dubbed as “weed-out” courses, consisting predominantly of traditional lecture-based passive education and cookbook-type laboratory courses, are blamed for a lot of these challenges, as students feel disappointed, bored, unsupported, overwhelmed, or intimidated in these courses (Tobias, 1991; Gainen, 1995; Seymour and Hewitt, 1997; Brainard and Carlin, 1998;

Graham Hatfull, *Monitoring Editor*
Submitted Dec 29, 2017; Revised Jul 5, 2018; Accepted Jul 11, 2018
CBE Life Sci Educ December 1, 2018 17:ar53
DOI:10.1187/cbe.17-11-0248

Potential conflict of interest: Authors T.R.-P. and L.X. assessed the effectiveness of the novel research experience for this paper and were also members of the design and administration team of the research experience. Deliberate efforts were made to ensure unbiased evaluation of outcomes. No commercially viable product was developed, and no promotion of a particular product to the exclusion of other similar products is involved.

*Address correspondence to: Vincent M. Cassone (vincent.cassone@uky.edu).

© 2018 T. Rodrigo-Peiris *et al.* CBE—Life Sciences Education © 2018 The American Society for Cell Biology. This article is distributed by The American Society for Cell Biology under license from the author(s). It is available to the public under an Attribution–Noncommercial–Share Alike 3.0 Unported Creative Commons License (http://creativecommons.org/licenses/by-nc-sa/3.0).
“ASCB®” and “The American Society for Cell Biology®” are registered trademarks of The American Society for Cell Biology.

American Association for the Advancement of Science [AAAS], 2011; PCAST, 2012). Despite aptitude and interest in STEM, some students underperform in these introductory STEM courses, and many others are dissatisfied with their grades, which negatively affect their science-related self-efficacy and cause switches to academic majors that they view as less challenging (Tobias, 1991; Gainen, 1995; Seymour and Hewitt, 1997; Brainard and Carlin, 1998; PCAST, 2012; Chen, 2013; Dagley *et al.*, 2016).

These concerns have urged expert working groups (e.g., Association of American Medical Colleges, 2009; AAAS, 2011; PCAST, 2012) to recommend urgent transformation of curricula and instructional methods to enhance student outcomes, and thereby better prepare them for 21st-century STEM workplace needs. Toward this end, AAAS (2011) has recommended six core competencies for biology undergraduate education (i.e., ability to apply the process of science, quantitative reasoning, modeling and simulation, understanding the interdisciplinary nature of science, communication and collaboration with other disciplines, and understanding the relationship between science and society) and urged biology educators to design their undergraduate curricula to target effective development of these competencies. Undergraduate research has been identified as a highly effective, evidence-based, active-learning method to address the challenges of STEM higher education across diverse students and has been recommended for the early undergraduate years (Russell *et al.*, 2007; Lopatto, 2010; Sadler *et al.*, 2010; AAAS, 2011; PCAST, 2012).

Two main designs of undergraduate research experiences have been used at the college level. In the longer-standing apprenticeship-based traditional undergraduate research experiences (ATUREs; elaborated in Hunter *et al.*, 2007; Russell *et al.*, 2007; Laursen *et al.*, 2010; Sadler *et al.*, 2010), the student conducts an independent research project in the research laboratory of a faculty member (i.e., research advisor) as part of the research advisor's overall research program and research laboratory group. As such, ATURE students typically receive one-on-one mentorship from the research advisor or his/her designated researcher from the research laboratory. Though highly beneficial to the undergraduates, ATUREs are generally resource intensive in terms of research expenses and faculty time per student, thereby limiting the number of opportunities available to students (Hunter *et al.*, 2007; Desai *et al.*, 2008; Wei and Woodin, 2011; Auchincloss *et al.*, 2014; Harvey *et al.*, 2014; Rodenbusch *et al.*, 2016). Arguably, ATUREs also require a considerable degree of preparedness and/or ability as well as self-confidence on the part of the student to meet the expectations of this rigorous and individualized experience (Hunter *et al.*, 2007; Weaver *et al.*, 2008; Junge *et al.*, 2010; Rowland *et al.*, 2012). Therefore, ATUREs are mostly pursued and secured by self-selected, science-committed, higher achievers, particularly in their later undergraduate years, who are preferentially recruited by the research laboratories (Hunter *et al.*, 2007; Russell *et al.*, 2007; Sadler *et al.*, 2010; Rowland *et al.*, 2012; Linn *et al.*, 2015; Rodenbusch *et al.*, 2016).

To provide the benefits of undergraduate research to a larger number of students at a lower per-student cost, particularly in the early undergraduate years, course-based undergraduate research experiences (CUREs) have become increasingly popular (elaborated in Wei and Woodin, 2011; Auchincloss *et al.*, 2014; Corwin *et al.*, 2015; Rodenbusch *et al.*, 2016). In CUREs, open-ended explorations of research questions of interest to the scientific community are pursued in a regular course setting, usually in an instructional laboratory space with a class of peer students, during scheduled course meeting times. Taking place within a structured and guided instructional environment, CUREs are generally designed so that all students who meet prerequisites to enroll in the course are provided the tools to succeed, which could be particularly beneficial for freshmen, novice researchers, and students with lower levels of preparedness (Nadelson *et al.*, 2010; Brownell and Kloser, 2015; Rodenbusch *et al.*, 2016). CUREs are diverse in their design and implementation. The more standardized CUREs that are broadly adopted across institutions include the Science Education Alliance–Phage Hunters (Hatfull *et al.*, 2006; Jordan *et al.*, 2014), Genomics Education Partnership (Shaffer *et al.*, 2010), and Small World Initiative (Barral *et al.*, 2014), henceforth referred to as "broad-based CUREs." In broad-based CUREs, faculty are generally recruited to facilitate the research course as instructors and contribute course-generated data to a networked results database, but they are not necessarily part of the original research team with intrinsic interest and professional research expertise in the research content (Wei and Woodin, 2011; Shortlidge *et al.*, 2016). Therefore, the students do not typically encounter interactions with the actual professional researchers. CUREs have also been developed and offered by institutions by incorporating faculty research projects or research interests into the laboratory courses (henceforth referred to as "local CUREs," as per Rodenbusch *et al.* [2016]). Research instructional guidance in local CUREs ranges from those in which the students are predominantly guided by course instructors, with the CURE collaborating with research professors (e.g., Bascom-Slack *et al.*, 2012; Harvey *et al.*, 2014), to those that are guided by the research professors serving as instructors, potentially with help from the research laboratory members (e.g., Nadelson *et al.*, 2010; Miller *et al.*, 2013).

In ATUREs, a supportive environment is provided for the student by a research culture–based tiered support system that includes the research advisor and other researchers in the research laboratory such as research scientists, postdoctoral researchers, graduate students, and other undergraduate researchers establishing interactive mentor–mentee and research collaborative relationships with the student (AAAS, 2011; Auchincloss *et al.*, 2014; Rodenbusch *et al.*, 2016). In CUREs, on the other hand, the instructor, instructional assistants, and the peers provide an instructional culture–based supportive environment to the student, as in other courses (Hatfull *et al.*, 2006; Rodenbusch *et al.*, 2016). Teacher–student type relationships with the instructors and their instructional helpers facilitate a learning-focused environment for the CURE participants to enhance knowledge and skills related to the research process via a "situated learning" experience (i.e., learning situated in an authentic activity) and provide vital feedback on how to improve (Lave and Wenger, 1991; Corwin *et al.*, 2015). To support learning for a class of students simultaneously, CUREs frequently incorporate explicit instructional approaches targeting specific learning outcomes, such as instructor explanations, lectures, demonstrations, discussions, assigned reading and writing, and team-based assignments (Shaffer *et al.*, 2010; Ditty *et al.*, 2013; Miller *et al.*, 2013; Harvey *et al.*, 2014;

Kowalski *et al.*, 2016), while ATUREs predominantly employ an implicit instructional approach to facilitate learning outcomes through research participation and associated experiences (Sadler *et al.*, 2010; Linn *et al.*, 2015). Therefore, facilitation of conceptual understanding is critiqued as more successful in CUREs compared with ATUREs due to such instructional support (Linn *et al.*, 2015), and superior learning of the nature of science has been reported when explicit instruction was integrated into an ATURE-like research experience that provided a hybrid explicit–implicit instructional experience (Schwartz *et al.*, 2004; Sadler *et al.*, 2010).

Assessing "pinnacle" outcomes of research programs (i.e., indicators of effectiveness important for stakeholders or to continue a program, such as persistence in science; Urban and Trochim, 2009; Corwin *et al.*, 2015) is particularly challenging due to the long-term nature of these outcomes and confounding secondary variables that also aggregate with time (Corwin *et al.*, 2015; Rodenbusch *et al.*, 2016). Also, evaluating long-term outcomes alone does not provide insights as to how these outcomes are affected by the research experience (Laursen *et al.*, 2010; Rodenbusch *et al.*, 2016). To address these setbacks, Corwin *et al.* (2015) used a systems approach to develop a pilot, multipoint, cause–effect model for one of the pinnacle outcomes (i.e., persistence in science), using the reported activities and outcomes of both CURE- and ATURE-like research experiences, learning theory, and logical reasoning. This model provides a basic framework to methodically assess CUREs for short-term, medium-term, and long-term outcomes (henceforth referred to as the "Corwin model"; the Corwin model is illustrated in Figure 5 of Corwin *et al.* [2015]).

Supported by a 5-year grant from the Howard Hughes Medical Institute (HHMI), in the 2014–2015 academic year, we initiated a freshman success–targeted program called the STEMCats program at the University of Kentucky (UK), with the main goals of enhancing student retention in STEM majors, 21st-century STEM competency development, and STEM academic performance across all participating students. This paper details 1) the research experience that was designed to engage the STEMCats students (henceforth referred to as the "STEMCats Research Experience" [SRE]), emphasizing on its design features; 2) the outcomes toward the main STEMCats program goals, evaluated via institutional data analyses and supporting student perceptions for its first iteration in the Spring 2015 semester; and 3) future strategies based on experiences and outcomes of this iteration.

RATIONALE, OBJECTIVE, AND RESEARCH QUESTIONS

To facilitate the targeted outcomes for STEMCats students, we determined that a highly promising evidence-based approach would be to facilitate an "authentic research" experience that would provide them with a rich academic engagement (Lopatto, 2010; Wei and Woodin, 2011) within a "supportive environment" that fosters collaborative, interactive, and motivating relationships (Nadelson *et al.*, 2010; Eagan *et al.*, 2013). We were also compelled to consider three main practical aspects when designing the SRE: scale, student factors, and resources.

In terms of scale, our cohort size of 119 students in this first offering was large-scale, particularly for ATURE experiences. Further, the scale needs for the future would likely be larger, and therefore to adjust to future demands, our experimental research experience design would benefit from an ability for rapid implementation and scalability.

In terms of student factors, due to the STEMCats program's open-enrollment format, which requires no academic prerequisites or prior performance requirements, and recruiting advisors particularly encouraging underrepresented and disadvantaged students (e.g., first generation, ethnic minorities, Pell grant recipients, low precollege achievements) to enroll in the program, the student population was expected to be diverse, in contrast to programs in which self-selection and competition for limited opportunities favor predomination by higher-achieving, self-driven, and advantaged students. Because the program was open to any STEM-interested freshman irrespective of academic major, the SRE students were also expected to be diverse in terms of their majors. With a 4-year targeted graduation, STEM freshman schedules are tight, and as college-adjusting second-semester freshmen, STEMCats students would also find the course work intensive and demanding. New at UK, these freshmen would be still adjusting to the university surroundings, people, and procedures; therefore, a logistically simple, predesigned, and easily accessible experience presented to them as "a course" seemed most ideal.

In terms of resources, we needed to consider faculty availability and time commitments and expenses per student, particularly considering the scale of the SRE. Also, we needed to garner these resources, including faculty enthusiasm, in a manner that would facilitate the experience to be offered in a sustainable manner in the future.

As ATURE or CURE designs themselves did not appear to be the best choice to address these needs for our context, we designed the SRE as a hybrid research experience, embedding features of ATUREs and CUREs in a manner that we reasoned was suitable to collectively address these five factors: 1) an authentic research experience, 2) a supportive environment, 3) current and future needs for scale, 4) student characteristics and circumstances, and 5) availability and sustainability of institutional resources.

The objective of this research study was to evaluate the two main design features (authentic research and supportive environment) and the intended student outcomes (retention in a STEM major, 21st-century STEM competency development, and STEM academic performance) in the first iteration of the SRE.

The research questions we seek to address are

- To what extent did the students perceive the SRE as facilitating "authentic research" and a "supportive environment"?
- To what extent did the SRE impact student retention in a STEM major?
- To what extent did the SRE impact 21st-century STEM competency development?
- To what extent did the SRE impact student academic performance in STEM?

METHODS

SRE Design and Implementation

For authenticity of the research experience, we sought to provide features as close as possible to those of an ATURE, envisioning that incorporating students into a true professional practice would provide its benefits by virtue of participating in a real research operation, thereby circumventing potential shortfalls or compromises inherent in attempting to mimic professional research practice within a classroom (i.e., "participate" vs. "simulate" concept, as per Barab and Hay [2001]). We reasoned that research projects connected to an authentic research program designed and mentored by the faculty member heading the research program and conducted in a research laboratory itself would be the most ideal. Faculty to lead the SRE course sections were recruited from various STEM disciplines to provide variety in the research projects, and 18 research faculty and three research-involved instructional faculty offered 20 SRE research projects based on their research interests (Supplemental Table S1). Thus, we were able to provide a range of choices for research projects to align with student interests, compared with the few potential options available had we provided regular CUREs instead, although the choices were not as varied as in ATUREs. Each SRE research project was a semester-long, 1–credit hour course section that occurred during a scheduled time as per the course schedule; some were led by collaborating pairs or a trio of faculty members. Other researchers in these research laboratories had varying mentoring responsibilities and research collaborations with the SRE students, as determined by the research advisors.

We accommodated a group of students per SRE to meet the scale needs and resource limitations, including faculty time availability and cost per student, while also providing peer-related benefits for the freshman students. Most sections were able to accommodate a freshman group of up to 12 students within a budget of $1000 for research expenses provided by the grant. The course enrollments based on student preferences yielded one section of 13 students, 10 sections of seven to 12 students, six sections of three to four students and three sections of fewer than three students. While ATURE students are typically expected to proceed with high self-reliance (Weaver *et al.*, 2008; Rowland *et al.*, 2012), these small peer groups in SRE sections were expected to ease freshman assimilation into the research environment; reduce the stress of self-responsibility; and facilitate conducting research and learning through peer instruction, team work, and so on. The smaller peer group size was also intended to minimize crowding effects and facilitate developing closer relationships, as compared with offering regular CUREs to replace freshman laboratory courses at UK that may consist of ~25 students per section. For the research professor, SRE implementation thus would not require the same level of planning and preparation or the same extent of student guidance, mentoring, and counseling as would offering a regular CURE experience to a large class of students (Shortlidge *et al.*, 2016).

We embedded a course structure from CURE features into the SRE design for the benefit of the freshman students and to facilitate organization and management of SRE sections. Therefore, faculty members or teams independently designed SRE sections within the overarching research course syllabus and broad design guidelines provided by central course administration based on STEMCats program goals. The specific design of the SRE sections varied based on the practices of STEM disciplinary areas, faculty creativity, and so on. However, each research project engaged the students in the key elements of an authentic research experience, including open-ended exploration of a scientific question of broader relevance; literature review; developing hypotheses; designing, conducting, and iterating experiments; collecting, analyzing, and interpreting data; and collaboration and scientific communication (Auchincloss *et al.*, 2014; Brownell and Kloser, 2015; Corwin *et al.*, 2015). Explicit instructional activities such as short guided-inquiry exercises, instructor explanations, practice sessions for experimental techniques, peer discussions, and assigned readings were integrated into the learning experience as part of the research course plan to develop students' knowledge and skills in research; this was in addition to implicit learning through conducting the research project. Each research project design was reviewed to ensure compliance with the program needs. One SRE section is described in Swanson *et al.* (2016). Students were provided with the research project descriptions in advance of course registration and had the opportunity to resolve any questions or concerns before making their selections. With the availability of laboratory space, equipment, research supplements, expertise, and trained researchers for guidance from within the research laboratories, the research advisors were able to initiate the projects in a short time and in a cost-saving manner, which would also be advantageous for future potential expansions.

In this first pilot offering, the SRE constituted an elective credit hour in the students' degree map, and most students enrolled for an SRE section free of charge as an additional credit hour beyond the fixed tuition fee for a full-time course load. Course enrollment occurred as per regular course registration procedure at UK; therefore, STEMCats freshmen were able to register for the SRE "course" as they would for any other course, without having to proactively approach and negotiate ATURE experiences with research faculty, which is known to be particularly discouraging to students from disadvantaged backgrounds and those underprepared for college or underrepresented in STEM (Lundberg and Schreiner, 2004; Bangera and Brownell, 2014). A Blackboard course management system shell established for each SRE section provided research advisors with a platform to organize and manage their SRE sections' activities. A research poster presentation by each student group during the last week of the semester at UK's annual institutional undergraduate research forum served as the culminating experience to provide a long-range goal for the semester. In this supportive research forum that emphasizes student participation, the student groups were able to present what they achieved during the semester without being pressured to produce higher-order results. All SRE groups participated in the STEMCats poster competition during the research forum, with certificates awarded for the top three poster performances, including oral presentation and discussion of the poster content.

The 1–credit hour workload that accounted for no more than 3 hours of contact time per week, lack of formal exams, and pass/fail grading based on participation and effort were expected to provide relief to the freshman schedules, workloads, and grade-associated stress levels, as concerns over workloads and grades are known to encourage switches from STEM majors into non-STEM majors (Seymour and Hewitt, 1997; PCAST, 2012; Dagley *et al.*, 2016). The targeted research achievements were also determined at the discretion of the research advisor of each section and evolved throughout the semester, based on what each advisor deemed achievable by his or her student group. This flexibility was intended to provide opportunity for each section to tailor the experience to respective student needs and abilities, which is particularly valuable for freshmen from diverse levels of preparation and STEM majors. Each SRE section was provided help from an undergraduate instructional assistant (a high-performing junior or senior from a STEM major) or a graduate teaching assistant engaged by the faculty advisor for instructional assistance and near-peer mentoring as needed. Thus, for student support, we provided a combination of a research culture–based ATURE support system (i.e., a hierarchical expertise-based support system consisting of an advisor and other laboratory members; Lave and Wenger, 1991; Rodenbusch *et al.*, 2016) and an instructional culture–based CURE support system (i.e., teacher–student relationships with instructor and instructional assistant, peer support, guided-learning facilitation, explicit instructional activities) that we envisioned as particularly effective for our freshman student population.

SRE Participants and Control Group

The 119 students who joined the STEMCats program during their enrollment as freshmen in the Fall 2014 semester continued participating in the STEMCats program in the Spring 2015 semester and took part in the SRE. As of Spring 2015, they represented traditional STEM majors available at UK (student numbers are shown following majors: biology [83], chemistry [20], physics [1], and pre-engineering [2]), STEM-related majors (health sciences [8], psychology [1], and equine science and management [1]), a non-STEM major (international studies [1]), and the "undeclared major" status designated as "undergraduate studies" (4). The traditional STEM majors for this study were identified based on a classification used by UK's academic database management. SRE participants are henceforth also referred to as "STEMCats students" or "STEMCats."

For comparative evaluation of retention in a STEM major and academic performance in STEM courses using institutional data, the STEMCats sample was restricted to biology and chemistry majors for several reasons. Biology and chemistry majors predominated among the STEMCats students and were represented by sufficiently large sample sizes compared with other majors that were represented by fewer STEMCats. These disciplines also represent two main traditional STEM majors faced with the previously discussed STEM education- and workforce-related issues that the national policy and funding agencies, including the HHMI Sustaining Excellence grant that funded this study, intend to address. These being traditional STEM majors, also facilitated the evaluation of retention within a STEM major more accurately. Further, due to the curricular diversity of academic majors, STEM academic performance evaluations could be more reliably evaluated when limited to these two majors.

The demographic, socioeconomic, residency, and academic details of these 103 biology and chemistry major STEMCats are detailed in Table 1. These STEMCats consisted of 21.12 and 20.41% of the biology and chemistry second-semester freshmen at UK, respectively. Student data were obtained from the UK institutional database, and the classifications within each parameter and computations are according to the standard methods used by institutional academic database management. High school grade point average (GPA) weighted for Advanced Placement (AP) classes and credits is reported on a scale of 0.00–5.00. UK GPA weighted for credits is reported on a scale of 0.00–4.00. Earned credit hours at the end of UK first semester have been computed using courses completed with a "D" grade or better at UK.

For conducting comparative statistical analyses of student outcomes, a control group of non-STEMCats students was prepared as follows. All non-STEMCats students who matriculated as freshmen into the same academic majors as the STEMCats in Fall 2014 and persisted at UK in the Spring 2015 semester as a biology or chemistry major were identified from the UK institutional database (388 students). After processing the data to remove ambiguous institutional data and students with obviously mismatched criteria compared with the STEMCats, the resulting control group yielded 376 students. The demographic, socioeconomic, residency, and academic composition of this non-STEMCats control group and their statistical comparisons with the STEMCats group are detailed in Table 1. Two-tailed statistical tests (z test for proportions and independent samples, t test for means) were performed for the compositional comparisons. Hedges's g effect sizes for t tests (unequal sample sizes) were calculated as a measure of practical importance to supplement statistical significance results (Hedges, 1981; Maher *et al.*, 2013). As predicted due to open-enrollment and targeted recruitment of disadvantaged and underrepresented students (e.g., ethnic minorities, Pell Grant recipients, first generation), STEMCats consisted of a diverse student body, with statistically nonsignificant incoming preparation and UK first-semester performance levels compared with the control group (Table 1).

Institutional Data

A comparative evaluation of outcome variables with respect to retention in a STEM major and STEM academic performance was conducted between STEMCats and the control, based on UK institutional data.

Outcome Variables. To evaluate the effectiveness of the STEMCats program, we evaluated STEM retention outcomes at two time points (discrete variables) and for five STEM performance outcomes: STEM course enrollment, STEM credit enrollment, STEM course pass rate, STEM credits earned, and STEM GPA (continuous variables). Both a descriptive data comparison and a regression analysis were conducted for each of these outcome variables.

The two time points used for STEM retention analyses are 1) the beginning of the Fall 2015 semester (can be considered "freshman-year STEM retention," because it accounts for the students who completed freshman year and started the sophomore year in a STEM major) and 2) the end of the Spring 2016 semester (can be approximated to "sophomore-year STEM retention," because most students complete their academic year in the Spring semester [i.e., take the Summer semester off] and hence could be considered as accounting for the students who completed the sophomore year in a STEM major). Any student who remained at UK with a declared major among any traditional STEM major at these time points was considered as a positive outcome. For logistic regression, we dummy-coded outcome variables: 1 = a STEM major, 0 = not a STEM major. These two STEM retention outcome variables are henceforth referred to as 1) freshman-year STEM retention and 2) sophomore-year STEM retention.

The five STEM performance outcomes were evaluated at the end of the Spring 2016 semester, which meant the end of sophomore year for most students, except for those few who took summer classes. Each of these STEM performance

TABLE 1. Summary of composition and statistical comparisons between STEMCats and the control group

	STEMCats (n = 103)				Control (n = 376)				Statistical test outcome[a]		Effect size
Variable	N	%	M	SD	N	%	M	SD	z-score	t (df)	Hedges's g
Female	65	63.1			243	64.6			–0.28 (p = 0.779)		
Race/ethnicity											
White or Caucasian	60	58.3			272	72.3			–2.73**		
Hispanic or Latino	15	14.6			16	4.3			3.75**		
Black or African American	12	11.7			32	8.5			1.00 (p = 0.317)		
Asian	10	9.7			33	8.8			0.28 (p = 0.779)		
Multiracial	4	3.9			13	3.5			0.19 (p = 0.849)		
Unknown	2	1.9			10	2.7			–0.46 (p = 0.646)		
Pell Grant recipient	29	28.2			110	29.3			–0.22 (p = 0.826)		
First generation	27	26.2			85	22.6			0.76 (p = 0.447)		
Out of state	37	35.9			116	30.9			0.96 (p = 0.337)		
Academic major: Spring 2015											
Chemistry	20	19.4			74	19.7			–0.07 (p = 0.944)		
Biology	83	80.6			302	80.3			0.07 (p = 0.944)		
High school GPA (0.00–5.00)	103		3.76	0.53	376		3.87	0.50		1.95 (477) (p = 0.052)	–0.22
Math ACT[b]	97		25.91	4.28	350		26.31	4.58		0.77 (445) (p = 0.441)	–0.09
UK first-semester GPA (0.00–4.00)	103		3.01	0.95	376		3.03	0.92		0.19 (477) (p = 0.846)	–0.02
UK first-semester earned credits	103		28.82	17.46	376		28.16	17.35		0.34 (477) (p = 0.733)	0.04

[a] z test for proportions, and independent-samples t test for means.
[b] Lower sample sizes for Math ACT due to missing scores in the database.
**$p < 0.01$ (two-tailed).

outcomes was evaluated for lower-division and upper-division STEM courses separately, using common STEM courses taken by biology and chemistry majors irrespective of STEMCats status. The lower-division STEM courses used for the calculation are the 100-level introductory chemistry, biology, and math courses available for STEM majors, and the upper-division STEM courses used for the calculation are the chemistry, biology, and math courses at the 200-level or above available for STEM majors (Supplemental Table S2).

1. STEM course enrollment: Number of enrolled lower-division or upper-division STEM courses. Courses with grades of “A” through “E,” “W” (withdrawal), and “I” (incomplete) accounted for enrolled courses.
2. STEM credit enrollment: Number of enrolled credits for lower-division or upper-division STEM courses. Courses with grades of “A” through “E,” “W” (withdrawal), and “I” (incomplete) accounted for enrolled credits.
3. STEM course pass rate: Percentage of lower-division or upper-division STEM courses passed by each enrolled student by scoring a “D” grade or better.
4. STEM credits earned: Credits earned for the lower-division or upper-division STEM courses by scoring a “D” grade or better.
5. STEM GPA: Credit-weighted lower-division and upper-division STEM GPAs were calculated manually and ranged on a scale of 0.00–4.00. The points allocated for each grade were “A” = 4, “B” = 3, “C” = 2, “D” = 1, “E” = 0. Courses with “W” (withdrawal) and “I” (incomplete) grades were not included in the GPA calculation, as per UK GPA calculation format.

Outcomes Assessment by Descriptive Statistics. Two-tailed statistical tests (z test for proportions and independent-samples; t test for means) were performed for retention and performance outcomes evaluation by comparative descriptive analysis. Hedges's g effect sizes for t tests (unequal sample sizes) were calculated as a measure of practical importance to supplement statistical significance results (Hedges, 1981; Maher *et al.*, 2013).

Outcomes Assessment by Regression Analysis. Based on literature, data accessibility, and our hypotheses, a set of

10 covariate factors were used as control variables in the regression analyses to evaluate the retention and performance outcomes between STEMCats and the control sample, thereby accounting for compositional differences between the two groups shown in Table 1 that may affect the outcomes (Schafer and Kang, 2008). On average, women, STEM minorities, and students from lower socioeconomic backgrounds (e.g., low income, first generation) have been shown to be disadvantaged in terms of STEM success compared with their counterparts (Seymour and Hewitt, 1997; PCAST, 2012; Riegle-Crumb *et al.*, 2012; Eagan *et al.*, 2013; Corwin *et al.*, 2015; Rodenbusch *et al.*, 2016). Further, UK institutional data show that out-of-state students show higher college attrition rates than in-state students. Also, student success parameters show variation at UK based on the academic STEM major, including retention rates and course performance. Precollege preparation, such as high school GPA and Math ACT achievements are also known to be predictors of college STEM success (PCAST, 2012; Riegle-Crumb *et al.*, 2012; Eagan *et al.*, 2013; Wang, 2013). STEM course taking and performance have been shown as indicators for STEM persistence in college, particularly in the freshman year (Chen, 2013). Therefore, college first-semester academic performance, which likely represents a combined effect of precollege preparation and college adjustment, may underlie subsequent STEM success. These first-semester controls would also serve to mitigate confounding effects of pre-SRE experiences on SRE outcomes. Thus, the 10 factors used as controls in the regression analysis are discrete variables: gender, ethnicity, Pell Grant recipient status, first-generation status, in state/out of state status, academic major; and continuous variables: high school GPA, Math ACT, first-semester UK GPA, and first-semester UK earned credit hours. The students with lower propensity for STEM success based on any of these factors would be considered "at-risk" students in this study.

The STEMCats and control sample data used for the descriptive analysis were further processed to prepare for the regression analysis. Students were categorized as "STEM minorities" and "STEM non-minorities" to create two broad ethnic groups in the regression analysis, based on the norm that "Black or African American," "Hispanic or Latino," "American Indian or Alaskan Native," and "Native Hawaiian or Pacific Islander" are considered STEM minorities, while "White or Caucasian" and "Asian" are not considered as STEM minorities (henceforth referred to as "STEM non-minorities"; PCAST, 2012). Students with ethnicity designated as "unknown" or "multiracial" were removed, because they could not be categorized as STEM minorities or STEM nonminorities. Mahalanobis analysis (critical chi-squared at $\alpha = 0.001$) was conducted (SPSS Statistics v. 22.0) to identify multivariate outliers for the regression analysis, and two outliers identified in the control sample were removed. Also, students with entries missing for a control variable were eliminated by our regression settings. The characteristics of the resulting STEMCats ($n = 91$) and control sample ($n = 328$) used for the regression analyses are shown in Table 2. Missing entries for outcome variables determined the final sample sizes for STEM retention and performance analyses.

Addition of the 10 control variables in a stepwise manner in the regression analyses across the evaluated outcomes reduced the overall regression error in general, encouraging their inclusion in the final regression models. The SRE participation variable was dummy-coded as STEMCat = 1 and non-STEMCat (i.e., control) = 0. The discrete control variables were dummy-coded 1 and 0, as denoted in Table 3. Binary logistic regression for STEM-retention outcomes analysis and ordinary least squares (OLS) multiple regression for STEM performance outcomes analysis were performed with SPSS Statistics v. 22.0. Regression residual assumptions (i.e., normal distribution, homoscedasticity) were not violated, and multicollinearity issues were not detected.

Student Perceptions

Multiple student surveys were administered to evaluate student perceptions on diverse aspects of the SRE. Survey questions relevant to the analysis conducted in this study were identified from two surveys referred to as STEMCats survey 1 (administered 3 weeks before the conclusion of the SRE to all STEMCats present during a cocurricular event) and STEMCats survey 2 (administered to all STEMCats present at the completion of the poster-presentation event that culminated the research experience). STEMCats survey 1 evaluated students' perceptions on the two main design elements of the SRE (authentic research and supportive environment) and their perceived gains from the SRE toward learning and development. These Likert-scale survey items were set on a seven-point scale with the students rating perceived gains from "strongly disagree" to "strongly agree." In both surveys, open-ended questions were provided for students to write down descriptive comments. The three questions used for the analysis herein are "What do you like about the STEMCats research lab/experience?" from STEMCats survey 1 and "Please share with us one thing that you feel most intriguing in the STEMCats program, and briefly explain it" and "Please share with us what you learned the most in the STEMCats program, and briefly explain it" from STEMCats survey 2. The surveys were paper based and responses were collected anonymously. The response rates for STEMCats surveys 1 and 2 were 78.15 and 99.16%, respectively, and all usable responses were included in the analyses.

As a model developed using results from both CURE- and ATURE-like experiences, the Corwin model could arguably be used to assess student outcomes toward enhanced science/STEM persistence in research experiences not limited to CUREs. This model consists of three phases of evaluation. In the early-phase evaluation, which is predominated by short-term outcomes that are relatively easy to measure, the "cognitive/skill" development of the student as a result of engaging in core research activities is measurable. This includes scientific knowledge and skill gains that lead to the student's science self-efficacy development. In the middle-phase evaluation, the "social" development of the student due to collaborative and supportive activities in the research experience that lead to a sense of belonging of the student to a larger scientific/STEM community is measured. In the late-phase evaluation, which is predominated by long-term outcomes that are relatively complex to measure, outcomes that are facilitated through composite effects of both early and middle phases are measured. These include enhancement of science/STEM motivation and science/STEM

TABLE 2. Summary of composition and statistical comparisons of key predictors between STEMCats and the control group used for regression analyses

Variable	STEMCats ($n = 91$)				Control ($n = 328$)				Statistical test outcome[a]		Effect size Hedges's *g*
	N	%	M	SD	*N*	%	M	SD	*z*-score	*t* (*df*)	
Female	57	62.6			207	63.1			–0.09 $(p = 0.928)$		
Race/ethnicity											
STEM minority	24	26.4			44	13.4			2.98** $(p = 0.003)$		
STEM nonminority	67	73.6			284	86.6			–2.98** $(p = 0.003)$		
Pell Grant recipient	27	29.7			97	29.6			0.02 $(p = 0.984)$		
First generation	24	26.4			71	21.6			0.97 $(p = 0.332)$		
Out of state	30	33.0			86	26.2			1.28 $(p = 0.201)$		
Academic major: Spring 2015											
Chemistry	18	19.8			65	19.8			0.00 $(p = 1.000)$		
Biology	73	80.2			263	80.2			0.00 $(p = 1.000)$		
High school GPA (0.00–5.00)	91		3.79	0.53	328		3.90	0.48		1.89 (417) $(p = 0.060)$	–0.22
Math ACT	91		26.00	4.32	328		26.00	4.54		0.00 (417) $(p = 1.000)$	0.00
UK first-semester GPA (0.00–4.00)	91		3.08	0.92	328		3.04	0.92		0.37 (417) $(p = 0.714)$	0.04
UK first-semester earned credits	91		28.90	17.01	328		28.80	17.30		0.05 (417) $(p = 0.961)$	0.01

[a] *z* test for proportions, and independent samples *t* test for means.

$^{**}p < 0.01$ (two-tailed).

TABLE 3. Logistic regression predicting freshman-year and sophomore-year retention in a STEM major for STEMCats ($n = 90$) and control ($n = 328$) from biology and chemistry majors

	Freshman-year STEM retention			Sophomore-year STEM retention		
	Unstandardized coefficients		Odds ratio	Unstandardized coefficients		Odds ratio
	B	SE B	Exp (B)	B	SE B	Exp (B)
Constant	–2.496	1.362	0.082	–3.143**	1.193	0.043
STEMCat (1) vs. non-STEMCat (0)	0.493 $(p = 0.178)$	0.366	1.637	0.604* $(p = 0.049)$	0.308	1.830
High school GPA (weighted, out of 5)	0.244	0.359	1.277	–0.294	0.315	0.745
ACT Math	0.023	0.042	1.023	0.089*	0.037	1.093
Female (1) vs. male (0)	0.094	0.299	1.099	0.206	0.257	1.228
STEM minority (1) vs. STEM nonminority (0)	0.168	0.376	1.183	0.120	0.328	1.128
Out of state (1) vs. in state (0)	–0.488	0.313	0.614	–0.346	0.274	0.707
Pell Grant recipient (1) vs. nonrecipient (0)	–0.025	0.321	0.975	0.072	0.280	1.075
First generation (1) vs. not first generation (0)	–0.015	0.336	0.985	–0.133	0.289	0.875
Academic major at the beginning of research experience: chemistry (1) vs. biology (0)	0.552	0.404	1.737	0.247	0.320	1.280
UK first-semester GPA (weighted, out of 4)	0.693***	0.159	2.000	0.697***	0.153	2.008
UK first-semester earned credit hours	0.014	0.013	1.014	0.019	0.011	1.019
–2*log likelihood (−2LL)	349.077			440.842		
Nagelkerke R^2	20.3%			24.2%		
Chi-square	$\chi^2 = 56.185$, $df = 11$, $p < 0.001$			$\chi^2 = 78.976$, $df = 11$, $p < 0.001$		
Hosmer and Lameshow test	$p = 0.862$			$p = 0.317$		
Classification accuracy	83.5%			73.2%		

$^{*}p < 0.05$.

$^{**}p < 0.01$.

$^{***}p < 0.001$.

identity development, leading to enhanced science/STEM persistence. This model is particularly well-suited to use as a guiding framework to evaluate the SRE, in which authentic research was conducted in a collaborative environment that would thereby align with the cognitive/skill and social elements of Corwin model's early-phase and middle-phase, respectively. While the Corwin model remains to be tested for accuracy and can potentially be improved (Corwin *et al.*, 2015), we have assumed its validity for the purpose of this study based on the evidential, theoretical, and logical foundations used for its development, and thereby take a first step toward incorporating this model for SRE evaluation. Therefore, using this available framework to organize student-perceived gains from the SRE toward enhancing STEM persistence/retention, we identified eight Likert-scale survey items from STEMCats survey 1 that matched closely with the early-, middle-, and late-phase evaluations/outcomes of the Corwin model.

To evaluate gains from the SRE toward 21st-century STEM competency development based on student perceptions, we identified 10 Likert-scale survey items from STEMCats survey 1 that matched closely with the core competencies recommended in AAAS (2011). Although our STEMCats group also included non–biology majors, we were motivated to evaluate the competencies delineated in this report due to the general applicability of these core competencies to other STEM majors, the abundance of life sciences–related research projects among STEMCats research experiences (i.e., 17 out of 20 SRE sections), and the predominance of life sciences–related majors in the STEMCats cohort (i.e., 78.15%), who also formed 41.67% in the three physical sciences SRE sections. The chosen items from STEMCats survey 1 corresponded closely to three out of the six core competencies recommended in the AAAS (2011) report, which are the "Ability to apply the process of science," "Ability to communicate and collaborate with other disciplines," and "Ability to understand the relationship between science and society." The other three competencies remain to be evaluated in a future study. On the other hand, the student-perceived gains from the SRE toward improving STEM academic performance of the participants were evaluated using ratings to three Likert-scale survey items in the STEMCats survey 1 that are logically relatable to STEM course performance.

However, it should be noted that the single-item survey measurement for each of these tested constructs/outcomes in our student perceptions analysis remains a limitation. Thus, student perceptions should only be considered as supporting evidence for the institutional data, where available.

RESULTS

Authentic Research and Supportive Environment Features of the SRE

The majority of the students perceived that the authentic research and supportive environment features of the SRE were strong, based on responses on the seven-point Likert scale: 92.47% of the STEMCats answered "somewhat agree" to "strongly agree" (Likert 5–7) regarding the SRE's fulfillment of authentic research, while only 5.38% of the STEMCats answered "strongly disagree" to "somewhat disagree" (Likert 1–3), in response to the survey item in STEMCats survey 1 that stated "Research lab experience—Integrated you to an authentic research community," and 2.15% of the STEMCats answered "neither agree nor disagree" (Likert 4). For supportive environment, 89.01% of the STEMCats answered "somewhat agree" to "strongly agree" (Likert 5–7), while only 3.30% of the STEMCats answered "strongly disagree" to "somewhat disagree" (Likert 1–3), in response to the survey item in STEMCats survey 1 that stated "Research lab experience—Provided you a supportive environment," and 7.69% of the STEMCats answered "neither agree nor disagree" (Likert 4; Figure 1 and Supplemental Table S3). Student comments provided in response to the open-ended survey questions "What do you like about the STEMCats research lab/experience?" in STEMCats survey 1 and "Please share with us one thing that you feel most intriguing in the STEMCats program, and briefly explain it" in STEMCats survey 2 revealed that they recognized and appreciated the authentic nature of the research experience and that they appreciated the support from the members in their research environment, including research advisors and STEMCats group mates (Table 4).

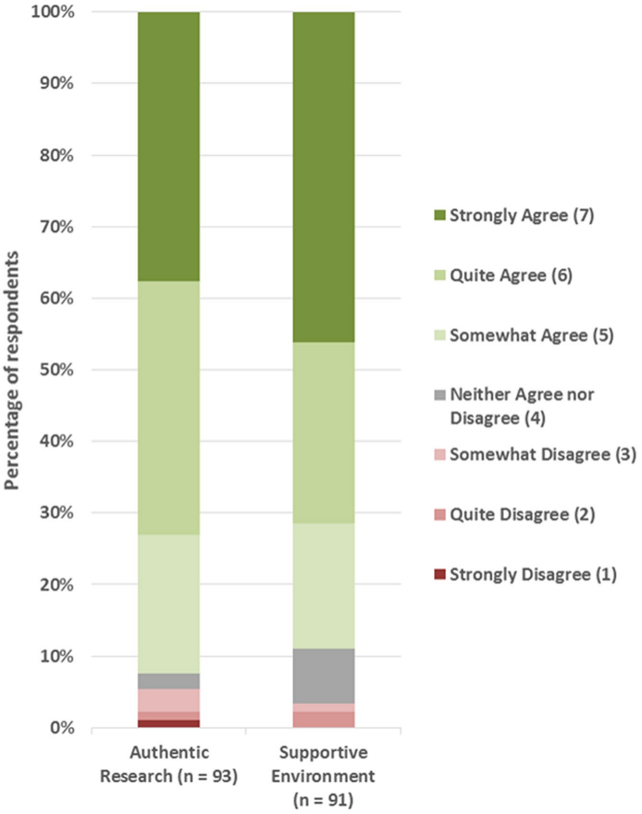

FIGURE 1. Student perceptions of "authentic research" and "supportive environment" features of the SRE. Percentages of respondents among SRE participants who answered "strongly disagree" to "strongly agree" on a seven-point Likert scale regarding the SRE's fulfillment of authentic research and supportive environment are shown. Student ratings on the Likert scale were in response to the survey items in STEMCats survey 1 that stated "Research lab experience—Integrated you to an authentic research community" and "Research lab experience —Provided you a supportive environment." Percentage of respondents by Likert category is available in Supplemental Table S3.

TABLE 4. Sample student comments regarding the two main design features of the SRE

Design features of the SRE	Student comments
Authentic research	"The opportunity to perform real research with real researchers as mentors." "I really enjoyed the research opportunity, it was neat to experience real research in a lab." "The research opportunity was a selling point for me. I had never had a chance to do real research in high school." "As freshm[e]n we are involved on a real research project and one that has been ongoing for years."
Supportive environment	"Everyone was very friendly[,] knowledgeable and willing to help." "The way undergrad and graduate students are able to interact with professors during research projects" "I liked working with a group." "The community and mentorship" "Whenever we ask [a] question, we get [a] real answer."

Effect of the SRE on STEM Retention

Comparative descriptive statistics (Supplemental Figure S1 and Supplemental Table S4) and correlation analysis (Supplemental Table S5) showed higher rates for freshman-year and sophomore-year STEM retention for students who participated in the SRE compared with the non-STEMCats control. However, these comparative gains were not statistically significant at $\alpha = 0.05$. Because descriptive data analysis does not control for the differences in student characteristics between the STEMCats and control that may affect the outcomes, it may be deficient in the outcomes comparison. Therefore, we conducted logistic regression controlling for secondary variables. Logistic regression results, after controlling for the 10 variables discussed previously, showed that STEMCats accomplished a 1.637 times higher STEM retention compared with non-STEMCats for freshman-year STEM retention; however, the B coefficient was not statistically significant ($p = 0.178$). For sophomore-year STEM retention (i.e., cumulative outcome of both freshman and sophomore years of retention), STEMCats showed a 1.830 times higher STEM retention compared with non-STEMCats, with a statistically significant B coefficient ($p = 0.049$). Other significant predictors of retention in this analysis were UK first-semester GPA for both freshman-year and sophomore-year STEM retention ($p < 0.001$) and ACT math score for sophomore-year STEM retention ($p = 0.016$; Table 3).

Supporting the positive outcomes from regression analysis, the majority of students who participated in the SRE perceived that their gains were high with respect to enhancing STEM persistence based on early-, middle-, and late-phase outcomes of the Corwin model (Figure 2 and Supplemental Table S3). On the Likert scale of 1–7, 86.67, 88.89, and 91.21% of the STEMCats answered "somewhat agree" to "strongly agree" (Likert 5–7) to the three survey items selected as relevant to assess Corwin model's early-phase outcomes: "Improved your scientific thinking," "Improved your science/STEM knowledge," and "Improved your experimentation skills," respectively. In contrast, only 4.44, 2.22, and 4.40% of the STEMCats answered "strongly disagree" to "somewhat disagree" (Likert 1–3) to these survey items, respectively. As per the Corwin model, these scientific knowledge and skill development outcomes should lead to improving the students' "Science/STEM self-efficacy."

Also, on the Likert scale of 1–7, 88.04, 94.57, and 86.81% of the STEMCats answered "somewhat agree" to "strongly agree" (Likert 5–7) to the three survey items selected as relevant to assess the Corwin model's middle-phase outcomes: "Improved

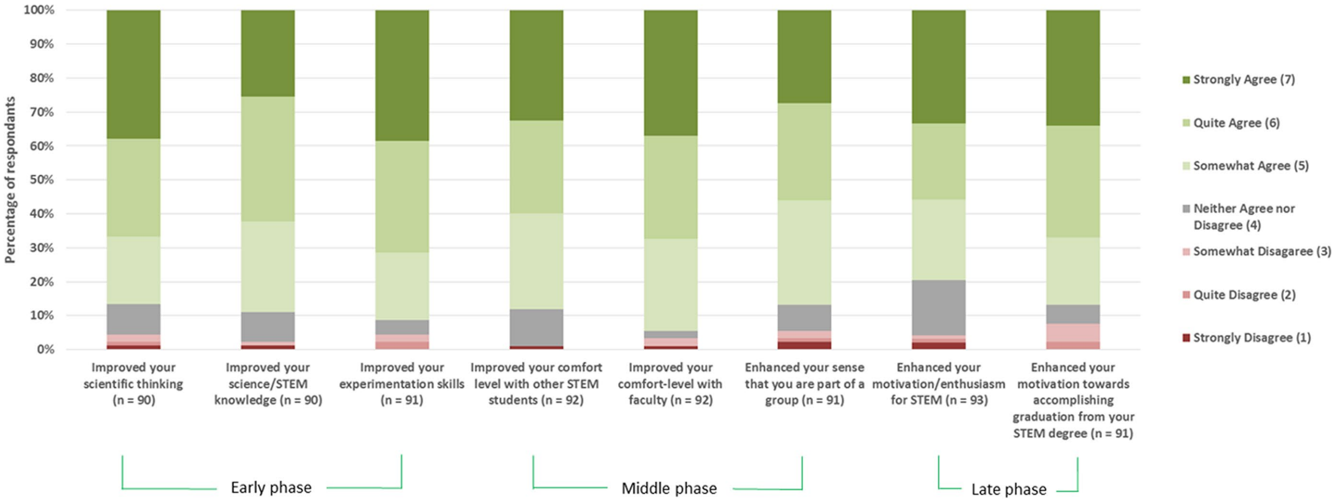

FIGURE 2. Student perceptions of STEM retention outcomes. Percentages of respondents among SRE participants who answered "strongly disagree" to "strongly agree" on a seven-point Likert scale regarding selected perceived gains toward STEM retention as per the Corwin model as a result of participating in the SRE are shown. These survey items from STEMCats survey 1 are categorized according to the corresponding evaluation phase (i.e., early, middle, late) of the Corwin model. Percentage of respondents by Likert category is available in Supplemental Table S3.

your comfort level with other STEM students," "Improved your comfort-level with faculty," and "Enhanced your sense that you are part of a group," respectively. In contrast, only 1.09, 3.26, and 5.49% of the STEMCats answered "strongly disagree" to "somewhat disagree" (Likert 1–3) to these survey items, respectively (Figure 2 and Supplemental Table S3). The third outcome, "Enhanced your sense that you are part of a group," we reasoned, closely corresponds with the Corwin model's middle-phase "hub" of "Sense of belonging to a larger community." A hub is a highly connected diagnostic outcome (Urban and Trochim, 2009; Corwin *et al.*, 2015).

Further, on the Likert scale of 1–7, 79.57 and 86.81% of the STEMCats answered "somewhat agree" to "strongly agree" (Likert 5–7) to the two survey items selected as relevant to assess Corwin model's late-phase outcomes: "Enhanced your motivation/enthusiasm for STEM" and "Enhanced your motivation towards accomplishing graduation from your STEM degree," respectively. In contrast, only 4.30 and 7.69% of the STEMCats answered "strongly disagree" to "somewhat disagree" (Likert 1–3) to these survey items, respectively (Figure 2 and Supplemental Table S3). The latter outcome, "Enhanced your motivation towards accomplishing graduation from your STEM degree," we reasoned, closely corresponds with the Corwin model's "pinnacle outcome" of "persistence in science." The reasoning used in selecting the STEMCats survey items with respect to Corwin model's evaluation phases, activities, and outcomes is elaborated in Supplemental Table S6.

Effect of the SRE on 21st-Century STEM Competencies

The majority of STEMCats perceived that their gains from the SRE were high toward enhancing the three evaluated 21st-century competencies from AAAS (2011): "ability to apply the process of science," "ability to communicate and collaborate with other disciplines," and "ability to understand the relationship between science and society" (Figure 3 and Supplemental Table S3). On the Likert scale of 1–7, 86.67, 90.00, 85.56, and 91.21% of the STEMCats answered "somewhat agree" to "strongly agree" (Likert 5–7) to the four survey items selected as relevant for assessing gains toward "ability to apply the process of science": "scientific thinking," "critical thinking," "trouble-shooting skills," and "experimentation skills," respectively. In contrast, only 4.44, 4.44, 4.44, and 4.40% of the STEMCats answered "strongly disagree" to "somewhat disagree" (Likert 1–3) to these survey items, respectively. Also, on the Likert scale of 1–7, 90.11, 88.76, and 88.76% of the STEMCats answered "somewhat agree" to "strongly agree" (Likert 5–7) to the three survey items selected as relevant for assessing gains toward "ability to communicate and collaborate with other disciplines': "knowledge in scientific communication," "teamwork skills," and "comfort level to work with colleagues from different academic backgrounds (e.g., different majors)." In contrast, only 3.30, 4.49, and 3.37% of the STEMCats answered "strongly disagree" to "somewhat disagree" (Likert 1–3) to these survey items, respectively. Further, on the Likert scale of 1–7, 86.67 and 87.91% of the STEMCats answered "somewhat agree" to "strongly agree" (Likert 5–7) to the two survey items selected as relevant for assessing gains toward "ability to understand the relationship between science": "improved your sense that science is connected to human lives" and "improved your sense that science is important to resolve real world issues," respectively. In contrast, only 1.11 and 3.30% of the STEMCats answered "strongly disagree" to "somewhat disagree" (Likert 1–3) to these survey items, respectively. Student comments further revealed that STEMCats gained these competencies and provided elaboration on their specific gains (Table 5). These comments were provided in response to

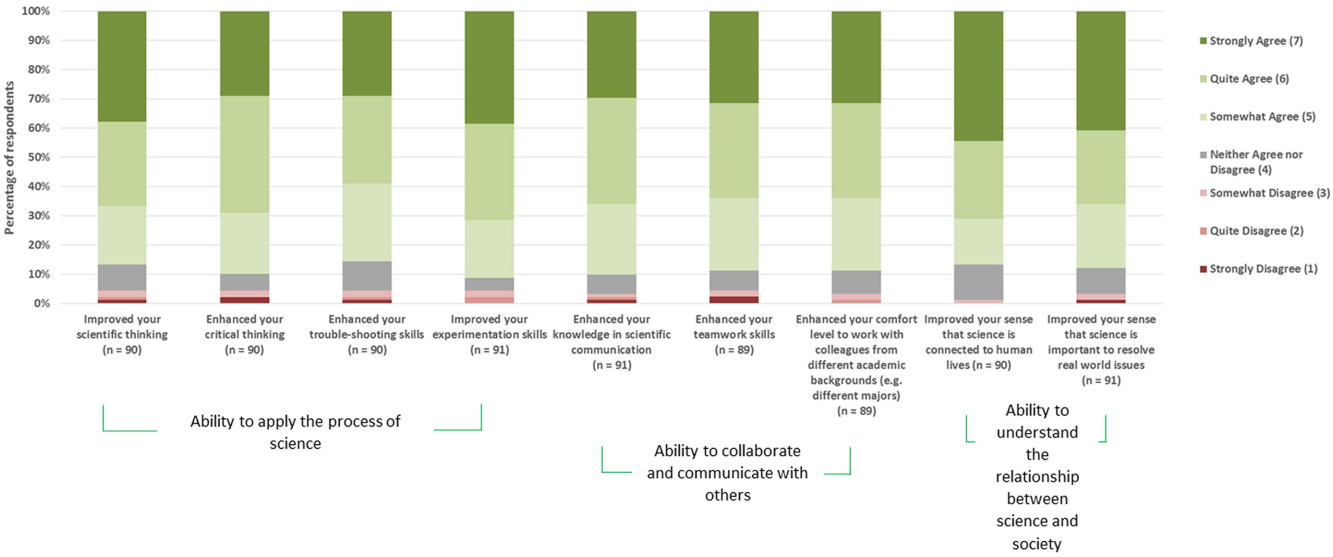

FIGURE 3. Student perceptions of STEM competency development outcomes. Percentages of respondents among SRE participants who answered "strongly disagree" to "strongly agree" on a seven-point Likert scale regarding selected perceived gains toward core STEM competency development as per AAAS (2011) as a result of participation in the SRE are shown. These survey items from STEMCats survey 1 are categorized according to the corresponding core competencies as specified in AAAS (2011; i.e., ability to apply the process of science, ability to collaborate and communicate with others, and ability to understand the relationship between science and society). Percentage of respondents by Likert category is available in Supplemental Table S3.

TABLE 5. Sample student comments regarding perceived gains from the SRE toward STEM competency development

AAAS (2011) core competencies	Student comments
Ability to apply the process of science	"I learned the process of research and [data] gathering and trying to explain data to answer questions." "Formulating research topics, analyzing data, writing posters & scientific reports" "How to think scientifically" "How to solve problems" "I learned how to use a microscope and other tools in the lab."
Ability to communicate and collaborate with other disciplines	"How to work in a pretty big group of people, and be considerate about everyone's input." "I learned how to work in teams." "I learned that communication is a large part of the STEM program with peers, mentors and professors." "Learned how to interact with more students" "Working with a group to achieve a goal."
Ability to understand the relationship between science and society	"We learned how [age] progression can effect autophagy in testis and ovary cells, thus causing infertility." "The research portion. It was nice to be able to see the real-life application." "I learned how to interpret calcium graphs of a heart cell and know how that affects the heart cell." "Although my group's research subjects were animals, we hope that the research itself can eventually be applied to human beings for medicinal and [other] purposes."

the open-ended survey questions "Please share with us what you learned the most in the STEMCats program, and briefly explain it" and "Please share with us one thing that you feel most intriguing in the STEMCats program, and briefly explain it" in STEMCats survey 2.

Effect of the SRE on STEM Academic Performance

In the comparative descriptive data analysis, STEM performance outcome variables: course enrollment, credit enrollment, course pass rate, earned credit hours for lower-division and upper-division STEM courses, and STEM GPA for lower-division STEM courses showed higher mean values for STEMCats compared with the control. However, statistical tests at $\alpha = 0.05$ only showed significant differences for course enrollment ($p = 0.009$, Hedges's g effect size 0.29), credit enrollment ($p = 0.007$, Hedges's g effect size 0.30), course pass rate ($p = 0.018$, Hedges's g effect size 0.27), and earned STEM credits ($p = 0.004$, Hedges's g effect size 0.33) for lower-division STEM courses. For upper-division STEM courses, no statistically significant differences were shown between STEMCats and the control group in this descriptive analysis (Supplemental Table S4 and Supplemental Figure S2). Correlation analysis showed statistically significant correlations with STEMCats participation only for lower-division STEM course enrollment, STEM credit enrollment, and STEM credits earned at $\alpha = 0.01$, and for lower-division STEM course pass rate at $\alpha = 0.05$ (Supplemental Table S5).

Because the descriptive data analysis does not control for covariables that may affect the outcomes, multiple linear regression analyses for the STEM performance outcomes, controlling for the 10 variables, were conducted with institutional data. The results revealed that, for lower-division STEM courses, STEMCats showed statistically significant higher achievements for course enrollment, credit enrollment, course pass rate, and earned credit hours ($p < 0.001$). However, the STEM GPA gain for lower-division STEM courses by STEMCats compared with the control ($p = 0.079$) was not statistically significant at $\alpha = 0.05$. For upper-division STEM courses, STEMCats showed statistically significant higher achievements for course enrollment, credit enrollment, and earned credit hours ($p < 0.01$). However, the gains for course pass rate ($p = 0.133$) and STEM GPA ($p = 0.055$) for upper-division STEM courses by STEMCats compared with the control were not statistically significant at $\alpha = 0.05$ (Table 6 and Supplemental Tables S7–S10).

Supporting the positive outcomes of the regression analysis, the majority of STEMCats perceived that their gains from the SRE were high toward the three selected gains relating to STEM academic performance: "motivation towards learning STEM," "science/STEM knowledge," and "understanding of scientific concepts" (Figure 4 and Supplemental Table S3). On the Likert scale of 1–7, 88.33, 88.89, and 91.21% of the STEMCats answered "somewhat agree" to "strongly agree" (Likert 5–7) to these survey items, respectively. In contrast, only 8.89, 2.22, and 2.20% of the STEMCats answered "strongly disagree" to "somewhat disagree" (Likert 1–3) to these survey items, respectively. The results of this study with respect to the four research questions that we addressed are summarized in Table 7.

DISCUSSION

The positive institutional data outcomes and limited yet corroborating student-perceived gains support that the SRE design appealed to the students and facilitated their progress toward the STEMCats program goals of retention in a STEM major, STEM competency development, and STEM academic performance. Owing to the diversity of student population in the STEMCats program, as indicated by equal or better representation of several at-risk and disadvantaged student categories among STEMCats compared with the control group, the SRE thus served a cross-section of students and not simply a student population more prone to STEM success.

Institutional data analysis that evaluated the "pinnacle outcome" of student retention in a STEM major showed a positive trend for STEMCats students compared with the non-STEMCats control immediately following the completion of the SRE (i.e., freshman-year STEM retention) in both the descriptive data analysis and the regression analysis, though not statistically significant. However, in the longer term (i.e., sophomore-year STEM retention: a year since the completion of the STEMCats

TABLE 6. Multiple linear regression predicting performance outcomes from lower-division and upper-division STEM courses, as of the end of sophomore year, for STEMCats (1) vs. control (0) from biology and chemistry majors[a]

	Lower-division STEM courses						Upper-division STEM courses					
	Beta	B	SE B	R^2	Adjusted R^2	Regression error	Beta	B	SE B	R^2	Adjusted R^2	Regression error
Course enrollment	0.193 $n_1 = 91$ $n_2 = 328$	1.135*** ($p < 0.001$)	0.252	28.6%	2.7%	2.078	0.078 $n_1 = 91$ $n_2 = 328$	0.425* ($p = 0.021$)	0.183	56.0%	54.8%	1.513
Credit enrollment	0.198 $n_1 = 91$ $n_2 = 328$	3.222*** ($p < 0.001$)	0.692	29.2%	27.3%	5.716	0.080 $n_1 = 91$ $n_2 = 328$	1.402* ($p = 0.019$)	0.595	54.7%	53.4%	4.911
Course pass rate	0.146 $n_1 = 86$ $n_2 = 311$	9.306*** ($p < 0.001$)	2.633	36.3%	34.5%	21.214	0.095 $n_1 = 58$ $n_2 = 192$	4.466 ($p = 0.133$)	2.961	10.8%	6.6%	19.224
Credits earned	0.206 $n_1 = 91$ $n_2 = 328$	3.769*** ($p < 0.001$)	0.738	36.7%	34.9%	6.095	0.080 $n_1 = 91$ $n_2 = 328$	1.391* ($p = 0.018$)	0.587	55.0%	53.8%	4.847
GPA	0.051 $n_1 = 88$ $n_2 = 312$	0.132 ($p = 0.079$)	0.075	69.3%	68.5%	0.607	0.097 $n_1 = 58$ $n_2 = 193$	0.212 ($p = 0.055$)	0.110	42.2%	39.6%	0.717

[a] The list of control variables is available in the *Methods* section and Supplemental Tables S7–S10.

n_1 = number of STEMCats.

n_2 = number of non-STEMCats.

*$p < 0.05$.

***$p < 0.001$.

research experience), STEM retention yielded a widened gap in the descriptive data comparison and a statistically significant positive result in the regression analysis favoring the STEMCats. Logistic regression results also corroborated that precollege math preparation (i.e., Math ACT score) and college performance (i.e., first-semester GPA) positively associate with college STEM retention (PCAST, 2012; Chen, 2013), while participation in the STEMCats program itself was a predictor for enhanced STEM retention, with a statistically significant effect on sophomore-year STEM retention when these covariate factors were controlled. It would be vital to conduct longitudinal analysis to evaluate whether these positive gains in STEM retention for STEMCats are sustained in subsequent years, leading to enhanced STEM graduation rates of the STEMCats. Enhanced STEM retention during progression through the academic years leading to higher production of graduates from STEM disciplines will contribute to ameliorating the projected deficit of college-educated STEM workforce, as recommended by PCAST (2012).

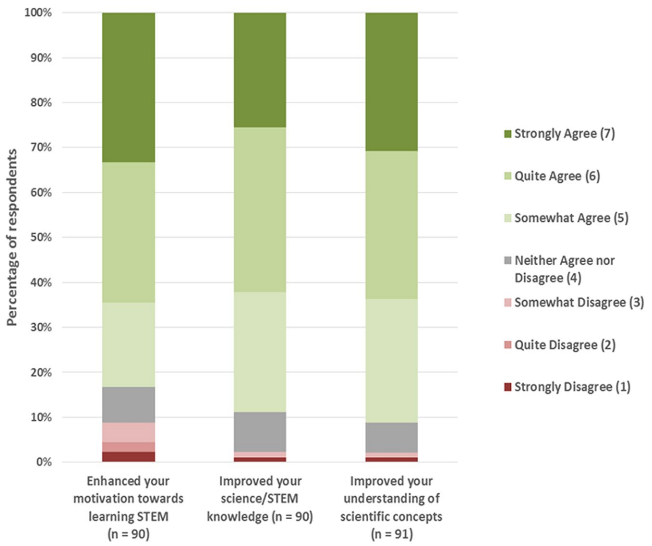

FIGURE 4. Student perceptions of STEM academic performance outcomes. Percentages of respondents among SRE participants who answered "strongly disagree" to "strongly agree" on a seven-point Likert scale regarding selected perceived gains toward STEM academic performance as a result of participation in the SRE are shown. These survey items were included in STEMCats survey 1. Percentage of respondents by Likert category is available in Supplemental Table S3.

Corroborating these enhanced retention trends observed for STEMCats in the institutional data, the high ratings assigned by the majority of STEMCats to the statements reflecting the early-, middle-, and late-phases of Corwin model support the ability to use the framework of the Corwin model to evaluate the impact of the SRE on student retention outcomes in a STEM major in a stepwise and insightful manner. According to the cause–effect relationships proposed in this model, the predominantly high student ratings received for the early- and middle-phase outcomes support that the SRE facilitated student "cognitive/skill" and "social" development, respectively, leading to the late-phase pinnacle outcome of enhanced science/STEM persistence. Based on this model, the perceived gains reported by STEMCats for the two late-phase outcomes in this cascade are noteworthy, because they each represented a medium-term outcome (i.e., "Enhanced your motivation/enthusiasm for STEM") and a long-term outcome (i.e., "Enhanced your motivation towards accomplishing graduation from your STEM degree") that were deemed challenging to achieve, yet were rated high by the majority of STEMCats within the one-semester, first iteration of the SRE. This suggests the SRE is an effective experience in enhancing

TABLE 7. Summary of results for the SRE, by research question

Evaluated feature or outcome	Regression	Descriptive data comparison	Student perceptions
Research question 1: Main design features			
Authentic research	—	—	92.47% rated from "somewhat agree (5)" to "strongly agree (7)."
Supportive environment	—	—	89.01% rated from "somewhat agree (5)" to "strongly agree (7)."
Research question 2: STEM retention			
Freshman year	Statistically nonsignificant positive outcome ($p = 0.178$)	Statistically nonsignificant positive outcome (z test, $p = 0.490$)	Positive gains in early-, middle-, and late-phase Corwin model outcomes toward STEM persistence, with more than 86% rated from "somewhat agree (5)" to "strongly agree (7)" for seven tested outcomes, and 79.57% for one tested outcome. Noteworthy that 86.81% rated from "somewhat agree (5)" to "strongly agree (7)" that the SRE enhanced their motivation toward graduation from STEM degree.
Sophomore year	Statistically significant increase at $\alpha = 0.05$	Statistically nonsignificant positive outcome (z test, $p = 0.091$)	
Research question 3: STEM competencies			
Ability to apply the process of science	—	—	More than 85% rated from "somewhat agree (5)" to "strongly agree (7)" for the tested four outcomes.
Ability to communicate and collaborate with other disciplines	—	—	More than 88% rated from "somewhat agree (5)" to "strongly agree (7)" for the tested three outcomes.
Ability to understand the relationship between science and society	—	—	More than 86% rated from "somewhat agree (5)" to "strongly agree (7)" for the tested two outcomes.
Research question 4: STEM academic performance			
Lower-division STEM course enrollment	Statistically significant increase at $\alpha = 0.001$	Statistically significant increase at $\alpha = 0.01$ (t test, effect size 0.29)	Positive gains for the three tested outcomes: "motivation towards learning STEM," "science/STEM knowledge," and "understanding of scientific concepts," with more than 83% rated from "somewhat agree (5)" to "strongly agree (7)."
Lower-division STEM credit enrollment	Statistically significant increase at $\alpha = 0.001$	Statistically significant increase at $\alpha = 0.01$ (t test, effect size 0.30)	
Lower-division STEM course pass rate	Statistically significant increase at $\alpha = 0.001$	Statistically significant increase at $\alpha = 0.05$ (t test, effect size 0.27)	
Lower-division STEM credits earned	Statistically significant increase at $\alpha = 0.001$	Statistically significant increase at $\alpha = 0.01$ (t test, effect size 0.33)	
Lower-division STEM GPA	Statistically nonsignificant positive outcome ($p = 0.079$)	Statistically nonsignificant slightly positive outcome ($p = 0.865$, t test, effect size 0.02)	
Upper-division STEM course enrollment	Statistically significant increase at $\alpha = 0.05$	Statistically nonsignificant positive outcome (t test, $p = 0.317$, effect size 0.11)	
Upper-division STEM credit enrollment	Statistically significant increase at $\alpha = 0.05$	Statistically nonsignificant positive outcome (t test, $p = 0.252$, effect size 0.13)	
Upper-division STEM course pass-rate	Statistically nonsignificant positive outcome ($p = 0.133$)	Statistically nonsignificant positive outcome (t test, $p = 0.365$, effect size 0.13)	
Upper-division STEM credits earned	Statistically significant increase at $\alpha = 0.05$	Statistically nonsignificant positive outcome (t test, $p = 0.264$, effect size 0.12)	
Upper-division STEM GPA	Statistically nonsignificant positive outcome ($p = 0.055$)	Statistically nonsignificant slightly negative outcome (t test, $p = 0.936$, effect size –0.01)	

the participants' STEM persistence. While the validity of the Corwin model was assumed for its pilot application in this study, regression analyses of student perceptions and objective data richly and closely reflecting the model's constructs, combined with time-point analyses, remain potential future directions for evaluating the Corwin model in a variety of undergraduate student research contexts, including the SRE.

The predominantly high ratings and comments reported by STEMCats regarding competency development supported that the SRE enhanced students' STEM competencies toward the three assessed AAAS (2011) competencies: science process skills, communication and collaborative competencies including working with colleagues from other academic disciplines, and understanding of the relatedness of science to the real world, which are vital for the rising STEM workforce. In further evaluations, objective testing via pre and post assignments spanning the SRE experience could add an insightful dimension to this analysis. Also, targeted efforts to enhance the AAAS (2011) competencies, as well as assessing the three competencies unassessed in this study, would be valuable in future iterations.

Descriptive data analysis and regression analysis supported the SRE's facilitation for enhanced STEM academic performance. Regression analysis yielded statistically significant results for STEMCats compared with the non-STEMCats control in all but one performance outcome for lower-division STEM courses (i.e., statistically significant gains in course enrollment, credit enrollment, course pass rate, and credits earned, but the gain in STEM GPA was marginally statistically insignificant at $p = 0.055$), and three outcomes for upper-division STEM courses (i.e., statistically significant gains in course enrollment, credit enrollment, and earned credit hours, but statistically insignificant gains in course pass rate and STEM GPA). Enhanced STEM enrollment rates for STEMCats suggest a higher undertaking of STEM course and credit loads together with their associated challenges, while the higher course pass rate for STEMCats students suggests greater success in overcoming these challenges. The higher number of STEM courses taken and greater success may have contributed toward higher STEM retention demonstrated by the STEMCats (Chen, 2013). The higher number of STEM credits earned by the midway point (i.e., end of sophomore year) of the targeted 4-year degree suggests that the STEMCats students were progressing at a faster rate toward accomplishing their STEM credits required for graduation compared with the control. Specifically designed pre and post assignments to evaluate STEM performance gains achieved by SRE participation could add another dimension to this analysis. The predominantly high ratings reported by STEMCats for perceived gains from the SRE toward enhanced motivation for learning STEM, science/STEM knowledge, and understanding of scientific concepts corroborated the results derived from institutional data, supporting SRE's facilitation for enhanced STEM academic performance.

In this study, several self-selection biases that may positively influence program outcome measures were mitigated in the descriptive outcomes analysis due to high recruitment of at-risk students to the STEMCats program. However, with controls applied for 10 prevalent secondary variables, we believe that the regression analysis presented is a more authentic representation of the outcomes compared with the descriptive data analysis. Yet there could be other influential variables that we may have omitted from the regression analysis. With regression conducted for institutional data, our study joins the few objective studies that have evaluated long-term outcomes of a research experience controlling for secondary variables, establishing causation of the research experience to outcomes (e.g., Junge *et al.*, 2010; Estrada *et al.*, 2011; Eagan *et al.*, 2013; Hernandez *et al.*, 2013; Rodenbusch *et al.*, 2016). Encouraged by the facilitation of SRE outcomes evaluation by the Corwin model and the defined biology competencies by AAAS (2011), future SRE, evaluations for STEM retention and competency development could be designed to be more closely aligned with these conceptual foundations and analyzed more comprehensively, perhaps including multiple survey items evaluating each construct within these conceptual frameworks and incorporating tested survey instruments (or adaptations thereof) for constructs when available. Variations between individual SRE sections in terms of design features, implementation, and student outcomes would also be vital to assess to gather further insights on the most effective sections and possible improvement of individual sections. Pooled data from several iterations and qualitative assessment methods may aid in these analyses, which would involve small sample sizes.

Given the positive implications from this evaluation of the SRE, which are consistent with the known positive impacts of undergraduate research on students, and the efficiency of offering the SRE design for a large number of students, further iterations within our institutional context and independent trials based on the SRE design in other institutions are recommended, with careful monitoring and continuous improvements. If the benefits are proven effective with the tests of repetition and scale, this experience could be offered to a broader audience of STEM freshmen. Consistent with the national recommendations to offer freshman research experiences for all students (AAAS, 2011; PCAST, 2012), SREs could serve as low-intensity research experiences for all STEM freshmen, and perhaps for non-STEM majors as well with necessary adaptations. However, under tight budgetary conditions limiting the scale, the SRE could be promoted predominantly for targeted students, such as at-risk students or other categories of students for whom the SRE may be particularly effective. The SRE design is well suited for research intensive institutions due to the availability of large numbers of research faculty and their laboratories. With the simplicity of this design, it also could be applied for freshman research in other types of institutions that are set up for authentic research, such as comprehensive and liberal arts institutions (Fairweather, 2005; Shortlidge *et al.*, 2016).

In addition to offering multiple sections of the most successful SREs and annexing additional new SRE sections to meet the demand for scale, the possibility of offering carefully designed local CUREs based on SREs could also be tested to accommodate more students at a lower cost, particularly under tight budgetary circumstances. Thus, local CUREs based on the most effective SRE sections could be piloted for a class of students (e.g., 20–30 students), led by the research advisor, and conducted in a quasi-research laboratory space. A method practiced by one SRE section, wherein the research sessions were held in an instructional laboratory near the research advisor's research laboratory, while the students also maintained constant contact with the advisor's research laboratory and its personnel, is an approach that could be tested to create a research laboratory feel for a CURE class to potentially enhance the authenticity of the research experience and also to provide a research culture–based support system to the participating students. Further, involving the advisor's research laboratory members as "research educators" in the CURE (i.e., a combined role of research mentor and instructor, as described in Rodenbusch *et al.* [2016]), may provide a hybrid research culture– and instructional culture–based supportive environment for CURE students. Careful evaluations, particularly in comparison with the SRE sections and ATURE students in these research laboratories, would provide insights regarding the effectiveness of these local CUREs.

The SRE design may also suit longer-term offerings of the research experiences over several semesters. With respect to ATUREs and CUREs, enhanced student outcomes have been reported with longer engagement in the research experience (Bauer and Bennett, 2003; Berkes, 2007; Russell *et al.*, 2007; Linn *et al.*, 2015; Rodenbusch *et al.*, 2016), with potentially enhanced benefits to the research advisor's research program as well. While we offered the SRE in the second semester of the freshman year to allow opportunity for freshmen to acquire a semester of college experience and STEM courses before conducting research in a real research laboratory, first-semester interventions for freshmen would be highly valuable to facilitate positive impacts from inception (Shapiro and Levine, 1999; Laufgraben, 2005; Dagley *et al.*, 2016). An earlier engagement with the SRE could be piloted via extending the SRE to the freshman first semester by forming SRE research groups at the beginning of the academic year itself. This first-semester course could be formulated as an "orientation to research" type, low-intensity, 1–credit hour course, with perhaps a weekly contact time of 1 hour as preparation toward the full-blown research project occurring in the next semester. Activities such as preparing solutions and other material for the upcoming research project may particularly bestow benefits related to active learning and project ownership to the students. A similar investigatory course preceding a research course series has shown positive impacts on student outcomes in Rodenbusch *et al.* (2016).

The per-student cost to offer the SRE in the current design averaged around $375, with $1000 for supplies, $3000 instructional fees for the faculty member as a teaching overload payment, and $500 payment for the undergraduate instructional assistant for an SRE section of 12 students. This per-student cost would generally be much lower compared with offering individual ATUREs, and also lower compared with the expenses mentioned in Rodenbusch *et al.* (2016) for their 3–credit hour CURE experience. The cost for disposables is also about half of that reported by Harvey *et al.* (2014) for their local CURE, and with no additional equipment costs in SRE sections for the start-up. Though the SRE is currently supported by grant funding, long-term sustainability in providing this experience to a large freshman population on a yearly basis would require financial planning. At an institutional level, positive student retention outcomes are expected to enable recovery of these expenses incurred due to inflow of tuition dollars in the subsequent years (PCAST, 2012; Rodenbusch *et al.*, 2016). The SRE instructional fees for the research faculty could be reduced or eliminated if the SRE could form a part of their regular teaching load and if the benefits of an SRE to a faculty member's research program were taken into consideration. Also, the undergraduate instructional assistants could be provided course credits, such as for leadership development or service learning, instead of a payment (Rodenbusch *et al.*, 2016).

While the main setbacks of the present study are that it is based on one iteration and a limited sample size that encumber a more definitive and detailed analysis, repeated future analysis, including the use of cumulative samples from subsequent cohorts, would help evaluate and elaborate these results. Also, to gain an insightful understanding of the effective features of the SRE design and implementation, as well as potential mechanisms by which the positive student outcomes were produced by the SRE, it would be valuable to conduct a detailed future study investigating these facets. Such a study could shed light on the underlying cause–effect relationships connecting design features with outcomes, as well as potential directions for future improvements. For example, following up on the predominantly high ratings by students for the main design features of the SRE, authentic research and supportive environment, it would be valuable to tease out the specific aspects of authentic research and support that were the most and least appealing to the participant students (i.e., specific factors underlying their ratings) and the degree to which the design features that were incorporated from ATURE and CURE designs were effective for this student group. On the other hand, the design features that were incorporated due to practical reasons, such as scale and limited resources, may have compromised the SRE's appeal or impact on the students. A detailed student perceptions' analysis on these aspects would be a first step toward revealing these insights. Further, it would be valuable to explore the impact of the SRE on different student populations, including making comparisons between the at-risk categories and their counterparts via analysis of both institutional data and student perceptions. Also, in the long term, the impact of the SRE experience on STEMCats students' career pathways could be tracked via a longitudinal study.

Amid the funding crisis for faculty-led research, the burden of costs for providing instructional laboratory experiences to undergraduates that are undereffective for the growing STEM workforce needs, and the financial challenges for curricular innovation (AAAS, 2011; PCAST, 2012; Edwards and Roy, 2017), the SRE appears a promising approach to facilitate student outcomes cost-effectively and at a substantial scale to address the STEM workforce demands, while also facilitating faculty-research, and systemic improvement.

ACKNOWLEDGMENTS

We thank Aaron Vaught for compiling institutional data, R. Joseph Waddington for suggestions on data analysis, and Margaret DiGirolamo for help with organizing data. We also thank the research advisors and their laboratory personnel for conducting SRE sections. This work was supported by an HHMI Sustaining Excellence-2014 grant (#52008116) awarded to the UK (V.M.C., principal investigator). The content and views expressed in this paper are the sole responsibility of the authors and do not necessarily represent the official views of the granting agency or the UK. This study was reviewed and approved by the UK Institutional Review Board (protocol 15-1089-P4S).

REFERENCES

American Association for the Advancement of Science. (2011). *Vision and change in undergraduate biology education: A call to action*. Washington, DC. Retrieved June 16, 2017, from https://visionandchange.org/files/2011/03/Revised-Vision-and-Change-Final-Report.pdf

Association of American Medical Colleges. (2009). *Scientific foundations for future physicians: Report of the AAMC–HHMI committee*. Washington, DC. Retrieved September 6, 2017, from www.aamc.org/download/271072/data/scientificfoundationsforfuturephysicians.pdf

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

Bangera, G., & Brownell, S. E. (2014). Course-based undergraduate research experiences can make scientific research more inclusive. *CBE—Life Sciences Education*, *13*(4), 602–606.

Barab, S. A., & Hay, K. E. (2001). Doing science at the elbows of experts: Issues related to the science apprenticeship camp. *Journal of Research in Science Teaching*, *38*(1), 70–102.

Barral, A. M., Makhluf, H., Soneral, P., & Gasper, B. (2014). Small world initiative: Crowdsourcing research of new antibiotics to enhance undergraduate biology teaching. *FASEB Journal*. Retrieved September 7, 2017, from www.fasebj.org/doi/abs/10.1096/fasebj.28.1_supplement.618.41

Bascom-Slack, C. A., Arnold, A. E., & Strobel, S. A. (2012). Plant microbiome and its products. *Science*, *338*(6106), 485–486.

Bauer, K. W., & Bennett, J. S. (2003). Alumni perceptions used to assess undergraduate research experience. *Journal of Higher Education*, *74*(2), 210–230.

Berkes, E. (2007). *Practicing biology: Undergraduate laboratory research, persistence in science, and the impact of self-efficacy beliefs (Doctoral Dissertation—Washington University in St. Louis)*. ProQuest Dissertations and Theses (3268009).

Brainard, S. G., & Carlin, L. (1998). A six-year longitudinal study of undergraduate women in engineering and science. *Journal of Engineering Education*, *87*(4), 369–375.

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.

Chen, X. (2013). STEM attrition: College students' paths into and out of STEM fields (NCES 2014-001). Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Retrieved February 2, 2018, from https://nces.ed.gov/pubs2014/2014001rev.pdf

Corwin, L. A., Graham, M. J., & Dolan, E. L. (2015). Modeling course-based undergraduate research experiences: An agenda for future research and evaluation. *CBE—Life Sciences Education*, *14*(1), es1.

Dagley, M., Georgiopoulos, M., Reece, A., & Young, C. (2016). Increasing retention and graduation rates through a STEM learning community. *Journal of College Student Retention: Research, Theory & Practice*, *18*(2), 167–182.

Desai, K. V., Gatson, S. N., Stiles, T. W., Stewart, R. H., Laine, G. A., & Quick, C. M. (2008). Integrating research and education at research-extensive universities with research-intensive communities. *Advances in Physiology Education*, *32*(2), 136–141.

Ditty, J. L., Williams, K. M., Keller, M. M., Chen, G. Y., Liu, X., & Parales, R. E. (2013). Integrating grant-funded research into the undergraduate biology curriculum using IMG-ACT. *Biochemistry and Molecular Biology Education*, *41*(1), 16–23.

Eagan, M. K., Hurtado, S., Chang, M. J., Garcia, G. A., Herrera, F. A., & Garibay, J. C. (2013). Making a difference in science education: The impact of undergraduate research programs. *American Educational Research Journal*, *50*(4), 683–713.

Edwards, M. A., & Roy, S. (2017). Academic research in the 21st century: Maintaining scientific integrity in a climate of perverse incentives and hypercompetition. *Environmental Engineering Science*, *34*(1), 51–61.

Estrada, M., Woodcock, A., Hernandez, P. R., & Schultz, P. W. (2011). Toward a model of social influence that explains minority student integration into the scientific community. *Journal of Educational Psychology*, *103*(1), 206–222.

Fairweather, J. S. (2005). Beyond the rhetoric: Trends in the relative value of teaching and research in faculty salaries. *Journal of Higher Education*, *76*(4), 401–422.

Gainen, J. (1995). Barriers to success in quantitative gatekeeper courses. *New Directions for Teaching and Learning*, *1995*(61), 5–14.

Harvey, P. A., Wall, C., Luckey, S. W., Langer, S., & Leinwand, L. A. (2014). The Python Project: A unique model for extending research opportunities to undergraduate students. *CBE—Life Sciences Education*, *13*(4), 698–710.

Hatfull, G. F., Pedulla, M. L., Jacobs-Sera, D., Cichon, P. M., Foley, A., Ford, M. E., … & Hendrix, R. W. (2006). Exploring the mycobacteriophage metaproteome: Phage genomics as an educational platform. *PLoS Genetics*, *2*(6), 0835–0847.

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational and Behavioral Statistics*, *6*(2), 107–128.

Hernandez, P. R., Schultz, P. W., Estrada, M., Woodcock, A., & Chance, R. C. (2013). Sustaining optimal motivation: A longitudinal analysis of interventions to broaden participation of underrepresented students in STEM. *Journal of Educational Psychology*, *105*(1), 89–107.

Hunter, A., Laursen, S. L., & Seymour, E. (2007). Becoming a scientist: The role of undergraduate research in students' cognitive, personal, and professional development. *Science Education*, *91*(1), 36–74.

Jordan, T. C., Burnett, S. H., Carson, S., Caruso, S. M., Clase, K., DeJong, R. J., … Hatfull, G. F. (2014). A broadly implementable research course for first-year undergraduate students. *mBio*, *5*(1)

Junge, B., Quiñones, C., Kakietek, J., Teodorescu, D., & Marsteller, P. (2010). Promoting undergraduate interest, preparedness, and professional pursuit in the sciences: An outcomes evaluation of the SURE program at Emory University. *CBE—Life Sciences Education*, *9*(2), 119–132.

Kowalski, J. R., Hoops, G. C., & Johnson, R. J. (2016). Implementation of a collaborative series of classroom-based undergraduate research experiences spanning chemical biology, biochemistry, and neurobiology. *CBE—Life Sciences Education*, *15*(4), ar55.

Laufgraben, J. L. (2005). Learning communities. In Upcraft, M. L., Gardner, J. N., Barefoot, B. O., & Associates (Eds.), *Challenging and supporting the first-year student: A handbook for improving the first year of college* (pp. 371–387). San Francisco: Jossey-Bass.

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

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. New York: Cambridge University Press.

Linn, M. C., Palmer, E., Baranger, A., Gerard, E., & Stone, E. (2015). Undergraduate research experiences: Impacts and opportunities. *Science*, *347*(6222), 1261757.

Lopatto, D. (2010). *Science in solution: The impact of undergraduate research on student learning*. Washington, DC: Council on Undergraduate Research.

Lundberg, C. A., & Schreiner, L. A. (2004). Quality and frequency of faculty-student interaction as predictors of learning: An analysis by student race/ethnicity. *Journal of College Student Development*, *45*(5), 549–565.

Maher, J. M., Markey, J. C., & Ebert-May, D. (2013). The other half of the story: Effect size analysis in quantitative research. *CBE—Life Sciences Education*, *12*(3), 345–351.

Miller, C. W., Hamel, J., Holmes, K. D., Helmey-Hartman, W. L., & Lopatto, D. (2013). Extending your research team: Learning benefits when a laboratory partners with a classroom. *BioScience*, *63*(9), 754–762.

Nadelson, L. S., Walters, L., & Waterman, J. (2010). Course-integrated undergraduate research experiences structured at different levels of inquiry. *Journal of STEM Education: Innovations & Research*, *11*(1), 27–44.

President's Council of Advisors on Science and Technology. (2012). *Engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics*. Washington, DC: U.S. Government Office of Science and Technology. Retrieved August 10, 2017, from https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/pcast-engage-to-excel-final_2-25-12.pdf

Riegle-Crumb, C., King, B., Grodsky, E., & Muller, C. (2012). The more things change, the more they stay the same? Prior achievement fails to explain gender inequality in entry into STEM college majors over time. *American Educational Research Journal*, *49*(6), 1048–1073.

Rodenbusch, S. E., Hernandez, P. R., Simmons, S. L., & Dolan, E. L. (2016). Early engagement in course-based research increases graduation rates and completion of science, engineering, and mathematics degrees. *CBE—Life Sciences Education*, *15*(2), ar20.

Rowland, S. L., Lawrie, G. A., Behrendorff, J. B. Y. H., & Gillam, E. M. J. (2012). Is the undergraduate research experience (URE) always best? The power of choice in a bifurcated practical stream for a large introductory biochemistry class. *Biochemistry and Molecular Biology Education*, *40*(1), 46–62.

Russell, S. H., Hancock, M. P., & McCullough, J. (2007). The pipeline: Benefits of undergraduate research experiences. *Science*, *316*(5824), 548–549.

Sadler, T. D., Burgin, S., McKinney, L., & Ponjuan, L. (2010). Learning science through research apprenticeships: A critical review of the literature. *Journal of Research in Science Teaching*, *47*(3), 235–256.

Schafer, J. L., & Kang, J. (2008). Average causal effects from nonrandomized studies: A practical guide and simulated example. *Psychological Methods*, *13*(4), 279–313.

Schwartz, R. S., Lederman, N. G., & Crawford, B. A. (2004). Developing views of nature of science in an authentic context: An explicit approach to bridging the gap between nature of science and scientific inquiry. *Science Education*, *88*(4), 610–645.

Seymour, E., & Hewitt, N. M. (1997). *Talking about Leaving: Why undergraduates leave the sciences*. Boulder, CO: Westview.

Shaffer, C. D., Alvarez, C., Bailey, C., Barnard, D., Bhalla, S., Chandrasekaran, C., … & Elgin, S. C. R. (2010). The Genomics Education Partnership: Successful integration of research into laboratory classes at a diverse group of undergraduate institutions. *CBE—Life Sciences Education*, *9*(1), 55–69.

Shapiro, N. S., & Levine, J. H. (1999). *Creating learning communities: A practical guide to winning support, organizing for change, and implementing programs*. San Francisco: Jossey-Bass.

Shortlidge, E. E., Bangera, G., & Brownell, S. E. (2016). Faculty perspectives on developing and teaching course-based undergraduate research experiences. *BioScience*, *66*(1), 54–62.

Swanson, H. I., Sarge, O. P., Rodrigo-Peiris, T., Xiang, L., & Cassone, V. M. (2016). Development of a course-based undergraduate research experience to introduce drug-receptor concepts. *Journal of Medical Education and Curricular Development*, *3*, 57–66.

Tobias, S. (1991). They're not dumb, they're different: Stalking the second tier. *American Journal of Physics*, *59*(12), 1155–1157.

Urban, J. B., & Trochim, W. (2009). The role of evaluation in research-practice integration: Working toward the "Golden Spike." *American Journal of Evaluation*, *30*(4), 538–553.

Wang, X. (2013). Why students choose STEM majors: Motivation, high school learning, and postsecondary context of support. *American Educational Research Journal*, *50*(5), 1081–1121.

Weaver, G. C., Russell, C. B., & Wink, D. J. (2008). Inquiry-based and research-based laboratory pedagogies in undergraduate science. *Nature Chemical Biology*, *4*(10), 577–580.

Wei, C. A., & Woodin, T. (2011). Undergraduate research experiences in biology: Alternatives to the apprenticeship model. *CBE—Life Sciences Education*, *10*(2), 123–131.