Adding Authenticity to Inquiry in a First-Year, Research-Based, Biology Laboratory Course

Jane L. Indorf,[†‡]* Joanna Weremijewicz,[†‡§] David P. Janos,[‡] and Michael S. Gaines[‡]

[‡]Department of Biology, University of Miami, Coral Gables, FL 33146; [§]Department of Biology, North Central College, Naperville, IL 60540

ABSTRACT

Course-based undergraduate research experiences (CUREs) are an effective way to integrate research into an undergraduate science curriculum and extend research experiences to a large, diverse group of early-career students. We developed a biology CURE at the University of Miami (UM) called the UM Authentic Research Laboratories (UMARL), in which groups of first-year students investigated novel questions and conducted projects of their own design related to the research themes of the faculty instructors. Herein, we describe the implementation and student outcomes of this long-running CURE. Using a national survey of student learning through research experiences in courses, we found that UMARL led to high student self-reported learning gains in research skills such as data analysis and science communication, as well as personal development skills such as self-confidence and self-efficacy. Our analysis of academic outcomes revealed that the odds of students who took UMARL engaging in individual research, graduating with a degree in science, technology, engineering, or mathematics (STEM) within 4 years, and graduating with honors were 1.5–1.7 times greater than the odds for a matched group of students from UM's traditional biology labs. The authenticity of UMARL may have fostered students' confidence that they can do real research, reinforcing their persistence in STEM.

INTRODUCTION

The best way to learn science is to do science. Yet science traditionally has been taught through prescriptive, "cookbook" laboratory exercises that do not expose students to scientific discovery. To improve student learning and persistence in science, many universities and colleges have replaced traditional laboratory courses with course-based undergraduate research experiences (CUREs). CUREs are a student-centered educational approach in which students conduct scientific research in the classroom. Auchincloss *et al.* (2014) defined CUREs as having all of the following five elements: 1) scientific practices, such as asking questions, proposing hypotheses, designing studies, collecting and analyzing data, and communicating results; 2) discovery of new knowledge; 3) relevant or meaningful work; 4) collaboration; and 5) iteration. CUREs involve students in many aspects of authentic research, because the research may be unscripted, the outcomes may be unknown, and the research can be highly relevant to the scientific community. Every genuinely authentic research experience entails involvement in the process of science and the creation of new knowledge (Table 1; Spell *et al.*, 2014; Rowland *et al.,* 2016).

Undergraduate laboratory course experiences lie along a continuum from prescribed laboratory exercises with a low level of inquiry to authentic CURES that give students much freedom (Table 1). Along this continuum are different levels of inquiry-based research experiences that involve students in the research process with increasing levels of independence and responsibility (National Research Council [NRC], 2000; Brown *et al.*, 2006; Table 1). In inquiry-based research, an instructor

TABLE 1. Laboratory course experiences in undergraduate science education range from structured inquiry-based labs to authentic research in the classroom that resembles traditional UREs outside the classroom

Laboratory course experience	Question type	Element[a]							
		Source of the question	Hypothesis formation	Data collection methods	Analysis of results	Iteration of scientific process	Type of equipment used	Data quality	Presentation of results
Structured inquiry (prescribed exercises)	Not novel	Given by instructor	Given by instructor	Given by instructor	Given by instructor	None	Basic	Confirmed by scientific literature	None
Guided inquiry	Not novel	Given by instructor	Open to student	Given by instructor	Given by instructor	None	Basic	Confirmed by scientific literature	None
Open inquiry	Not novel	Given by instructor	Open to student	Open to student	Open to student	May be included	Basic	Confirmed by scientific literature	None
Authentic research (UMARL)	Novel	Open to student	Open to student	Open to student	Open to student	Essential	Advanced	Potentially publishable	Scientific conference presentation
Traditional URE	Novel	Open to student/ Given by instructor	Open to student	Open to student/ Given by instructor	Open to student	May be included	Advanced	Potentially publishable	Scientific conference presentation

Authentic research-based courses such as UMARL engage students in many discovery-based elements of inquiry similar to traditional UREs, such as novel questions that are asked by the student, student experiment design, gathering of potentially publishable data, using sophisticated equipment, and presentation of results in a scientific conference presentation.
[a]The elements of inquiry per each level of laboratory course experience are gathered from findings by Buck *et al.* (2008), Weaver *et al.* (2008), Blanchard *et al.* (2010), Auchincloss *et al.* (2014), and Spell *et al.* (2014).

assigns the problem or question (Buck *et al.*, 2008; Weaver *et al.*, 2008) but gives students some independence to design and conduct their own investigations. In inquiry-based labs, results usually are not known to the students, but often are known to the instructor and scientific community (Dolan, 2016). As students are given increased freedom to ask their own research questions and devise their own investigations, the experience becomes increasingly *authentic*, with students engaging in the discovery process as scientists. The more freedom students have to make decisions about their projects, the more authentic the experience is. In CUREs, students conduct original research in that neither students nor instructors know the outcomes, which are relevant to current research within the scientific community. Furthermore, students may be provided the opportunity to use sophisticated lab equipment for their data collection (Weaver *et al.*, 2008), resulting in potentially publishable data. CUREs often culminate in public presentations of their findings (Weaver *et al.*, 2008; Table 1).

Because they are embedded within courses, CUREs can accommodate a greater number and more diverse group of students than traditional mentor–apprenticeship undergraduate research experiences (UREs), which often are limited by funding, space, and time (Desai *et al.*, 2008; Wei and Wooden, 2011; President's Council of Advisors on Science and Technology [PCAST], 2012; Bangera and Brownell, 2014). Typically, students selected for traditional UREs are high-achieving or upper-division science, technology, engineering, and mathematics (STEM) majors already motivated to succeed in their chosen field (Lopatto, 2007; Linn *et al.*, 2015). In contrast, CUREs can engage first- and second-year undergraduates of different abilities and backgrounds who may not yet have decided on a major or career pathway and are therefore vulnerable to not pursuing a STEM major, even if they are interested in STEM (Seymour and Hewitt, 2000). Hence, the broad inclusivity of CUREs may increase the number and diversity of students pursuing STEM degrees (Nagda *et al.*, 1998; Lopatto, 2004; Russell *et al.*, 2007; Bascom-Slack *et al.*, 2012; Hanauer *et al.*, 2012; Eagan *et al.*, 2013; Graham *et al.*, 2013).

CUREs provide many of the same benefits as traditional mentor–apprenticeship UREs (Lopatto, 2007; Shaffer *et al.*, 2010, 2014; Wei and Woodin, 2011; Auchincloss *et al.*, 2014; Table 1). Both experiences provide students a hands-on opportunity to participate in the discovery of novel information through participation in all scientific practices, such as reading and evaluating the scientific literature, selecting appropriate methods and experimental designs, collecting data, analyzing results, and working collaboratively (Corwin *et al.*, 2015). These activities have been shown to improve students' understanding of science and to increase their ability and confidence to conduct research, motivating them to learn science and self-identify as a scientist (Hunter *et al.*, 2007; Thiry and Laursen, 2009; Adedokun *et al.*, 2013; Corwin *et al.*, 2015). Thus, CUREs, like UREs, influence students' decisions to remain in STEM majors and, ultimately, to pursue STEM careers (Lopatto, 2004, 2007; Hunter *et al.*, 2007; Harrison *et al.*, 2011; Adedokun *et al.*, 2012; Graham *et al.*, 2013; Auchincloss *et al.*, 2014; Corwin *et al.*, 2015; Linn *et al.*, 2015).

CUREs have become a common way to provide authentic research experiences for students. For example, the multi-institution Science Education Alliance—Phage Hunters Advancing

Genomics and Evolutionary Science (SEA-PHAGES) program engages undergraduates in authentic research (Hanauer *et al.*, 2017) through discovery of novel bacteriophages from environmental samples (Jordan *et al.*, 2014). The multi-institutional Genomics Education Partnership (GEP) is a CURE integrated into genetics courses, in which students do authentic research in the field of genome evolution by improving *Drosophila* DNA sequence data and annotating chromosomes (Lopatto *et al.,* 2008; Shaffer *et al.*, 2010). While GEP and SEA-PHAGES are one and two semesters, respectively, the Freshman Research Initiative (FRI) at the University of Texas, Austin, is a three-semester research course series. FRI students start out by learning general research methods in the first semester. Following this, they move into a selected research stream. In the first research-stream course, they learn methods pertinent to their research topics and participate in guided inquiry-based projects. In the following course, students conduct independent authentic research and may have the freedom to ask their own research questions and design their own projects (Rodenbusch *et al.*, 2016; National Academies of Sciences, Engineering, and Medicine, 2017).

Although it is widely agreed that CUREs positively influence student learning and persistence in the sciences, many reports of CUREs have been descriptive (National Academies of Sciences, Engineering, and Medicine, 2017) with a lack of outcome data. Most studies have assessed small numbers of students, and few have had matched control groups, limiting conclusions about outcomes (Brownell *et al.*, 2012, 2013; Dolan, 2016; Rodenbusch *et al.*, 2016; National Academies of Sciences, Engineering, and Medicine, 2017). Evaluations of CUREs should control for students' backgrounds and academic preparation (Theobald and Freeman, 2014). Moreover, because student self-reporting does not always correlate well with actual learning (Falchikov and Boud, 1989; Shortlidge and Brownell, 2016), evaluation of CURE learning gains should be supported with objective measures of academic achievement (Linn *et al.*, 2015; National Academies of Sciences, Engineering, and Medicine, 2015). An exemplary study of FRI program participants conducted by Rodenbusch and colleagues (2016) has met these more rigorous requirements.

More than a decade ago, we developed and implemented a CURE in an introductory general biology lab course for first-year students at the University of Miami (UM). The UM Authentic Research Laboratories (the acronym "UMARL" reflects South Florida's prevalent limestone) courses imported faculty and their research into the classroom, where students conducted investigations of their own design within the realm of the faculty members' research topics. Students were given freedom to ask their own novel questions, to form hypotheses, and then to devise research that would generate data to test their hypotheses. Their results had the potential to contribute to the faculty instructors' research programs. Student-generated research questions and project designs conferred an authentic research experience.

The goals of this report are threefold. First, we describe the design and pedagogy of our UMARL course. Second, we compare student self-reported learning gains with national data. Third, using demographic and academic data of students who took UMARL over a span of 10 years and a matched comparison group from UM's traditional introductory biology labs, we evaluate outcomes, including participation in individual research, retention in STEM majors, time to graduation, and graduation with honors. We hypothesized that UMARL courses would have a positive influence on students' learning gains, because the unscripted research of UMARL engaged students in many of the same activities as traditional UREs, interesting them in subsequent research experiences, retaining them in STEM, and improving academic achievement (Tai *et al.*, 2006; Hansen and Birol, 2014; Olimpo *et al.*, 2016). We hypothesized that students in UMARL would report greater learning gains in research skills compared with all students responding to a national CURE survey, which potentially included respondents experiencing different authenticity. We also specifically predicted that UMARL students would be more likely to participate in individual research, graduate with a STEM degree, graduate with honors, and graduate within 4 years than a matched group of UM's traditional biology lab students.

METHODS

Course Design and Implementation

The goals of the authentic research UMARL course were to 1) lead students to the realization that the outcomes of research are not known in advance, hypotheses are not "proven," and things do not always go as planned; 2) expose students to what scientists do and how they think and to research as a career; and 3) engage research faculty in teaching first-year undergraduates, potentially fostering subsequent mentoring of undergraduates' individual research. The participation of research faculty also provided first-year undergraduates with an opportunity to interact individually with faculty, which is difficult in the large lecture setting of many introductory science courses and rarely happens in introductory labs, typically taught by graduate students.

In UMARL, four groups of five or six students devised and carried out their own investigations based on a faculty instructor's general area of research. Two sections, each with 20–24 students, ran concurrently, meeting for 3 hours per week. Halfway through the semester, faculty and their teaching teams interchanged their entire sections of students, such that students switched research themes. In each new research theme, students developed new novel questions, hypotheses, and experiments. After this midsemester switch, students stayed in their same groups, although they were free to form new research groups within their new lab sections in the following Spring semester. Thus, during the two-course, 1-year lab sequence, students were exposed to four different faculty members and their research areas. These authentic research labs were taken in place of the traditional two-semester introductory biology lab course sequence (i.e., students in UMARL earned introductory biology lab course credits).

Teaching Teams. Three-person teams consisting of a faculty member, a graduate teaching assistant, and an undergraduate peer facilitator (either paid a small stipend or awarded credit) led each section of four student groups. All three members of the teaching team were present throughout every lab session. The primary role of the faculty member was to bring his or her research program into the classroom and to provide guidance for a broad range of possible student investigations in a nonprescriptive manner. Faculty were funded investigators or

otherwise engaged in research. The graduate teaching assistant helped student groups design, conduct, and report their research (through PowerPoint presentations and scientific poster preparation) and assisted with grading. The undergraduate peer facilitator usually was an alumnus of UMARL who preferably had taken the lab from the same faculty instructor. The graduate teaching assistant and undergraduate peer facilitator fostered interactions among student groups, modeled collaboration with faculty, and underscored that the labs are student-centered, all of which can enhance student learning and academic performance (Otero *et al.*, 2006).

Research Themes. UMARL students were exposed to a diversity of research themes. Over 10 years, we had more than 20 different research themes led by as many different faculty, mostly from the biology department but also including faculty from UM's marine and medical schools (Table 2). Themes involved different numbers of independent and response variables, and those with a high product of independent and response variables offered students the most opportunity for creativity and control over the design of their investigations. For example, investigations of a glycoprotein called "glomalin," which is produced by ubiquitous arbuscular mycorrhizal fungi, involved measuring just a single dependent variable, the amount of protein in extracts of soil. Nevertheless, because glomalin can be found in almost every soil, students could test many different hypotheses regarding factors affecting its abundance. Another research theme was the health of soft corals (maintained in aquaria) assessed by three response variables (numbers of photosynthetic endosymbionts, chlorophyll concentration, and protein concentration of small tissue samples) after exposure to an environmental stressor (warming,

TABLE 2. Research themes that have been implemented in UMARL[a]

Research theme	Sample student research projects
Soft coral responses to pollutants	• Too much of a good thing? The effects of excess nutrients on anemone bleaching • The effect of salinity on zooxanthellae density and protein concentration of *Aiptasia pallida* • The effect of sunscreen pollution on the symbiotic relationship between *Aiptasia pallida* and their zooxanthellae • Oceans under ultraviolet radiation: metabolic changes in the *Aiptasia pallida*–Zooxanthellae symbiosis
Anticancer properties of plant extracts[b]	• Is *Caryota mitis* (fishtail palm) an anticarcinogen? • *Garcia nutans* (false tung oil tree) anticolon cancer clonogenicity assay • Anticancer activity in *Ptychosperma macarthurii* (Macarthur palm)
Glycoprotein production by mutualistic fungi[b]	• Do soil depth and root density have implications for glomalin concentration? • Abundance of mycorrhiza vs. concentration of glomalin • Glomalin in sun vs. glomalin in shade • Difference in glomalin levels beneath *Ficus benghalensis* and *Swietenia mahagonia*
Kinematic analysis of limbed locomotion in arthropods	• Effect of surface irregularities on stride length in scorpions and duty factor in terrestrial tarantulas • Effect of substrate differences on maximum sprinting speed in fiddler crabs
Respiration and metabolism	• Male cricket respiration • Absolutely breath taking: the effect of temperature on crickets' rate of oxygen consumption • Oxygen production rate of *Chara* (freshwater alga) with light • Oxygen consumption of crayfish in a dark environment
Population genetics of invasive and native species in South Florida	• Genetic variation among populations of pink shrimp (*Penaeus duorarum*) • Genetic variation among three populations of the brown anole (*Anolis sagrei*) • DNA sequence variation of the COX1 gene in tree snail (*Drymaeus dominicus*) populations • mtDNA Comparison of the Spanish sardine (*Sardinella aurita*) within South Florida
Stable isotopes (carbon-13 and nitrogen-15)	• Distinguishing between organic and inorganic produce by $^{15}N/^{14}N$ ratio • Carbon-13 and nitrogen-15 ratios of coffees from different latitudes • What's your beef? The use of stable isotopes to determine cattle diet • Apples to apples: adulteration of commercial juices with corn sugars
Gene expression	• The effects of lawn fertilizer on the metabolic gene expression of *Fundulus heteroclitus* • Motion in the ocean: the effect of water motion on the gene expression of *Fundulus heteroclitus* • The effects of petroleum on gene expression • The effects of alcohol on gene expression
Environmental effects on embryonic development	• Sea urchin fertilization: the effects of lawn fertilizer pollution on the motility and acrosome reaction of *Lytechinus pictus* sperm • Are they going to find me?: The effects of egg jelly on the fertilization of *Lytechinus pictus* • Does increasing insulin levels using Glipizide affect the development of *Danio rerio* embryos? • Nicotine hinders *Danio rerio* neurological development
Animal behavior	• Effects of apparent retinal image velocity in *Acheta domesticus* (house cricket) • The effect of parietal eye inhibition on the homing ability of *Anolis sagrei*, the brown anole • The relationship between variation in *Acheta domesticus* calling song inter-trill length and female responses • Effect of tail loss on *Anolis sagrei* running and jumping

[a]Student research projects for each theme took place over half a semester (6–7 wk). Details on some research themes and student projects can be found in Supplemental Table S1. Note that this table is not a complete list of all research topics implemented in the labs.
[b]These two themes involved extensive student use of the university's arboretum.

petrochemicals, fertilizer runoff, over-the-counter pharmaceuticals) limited primarily by students' imagination. Additional examples of research themes and suggestions for their implementation are available in Supplemental Table S1.

Course Progression and Deliverables. Each UMARL section followed the same general pattern and involved the same formative assessments (or deliverables). During the first class meeting, students were provided background information in a faculty lecture and were trained hands-on in analytical methods useful for investigating questions related to the research theme. The faculty instructor presented general examples of questions that student groups might pose but avoided being at all prescriptive. Each group of students then formulated their own research question, stated relevant hypotheses, and devised a protocol for testing them which the group presented orally at the start of the second lab. Additionally, students conducted a literature search for specific background information and potential methodologies, thus becoming acquainted with bibliographic search tools and interpreting scientific literature. Not only did student groups conduct investigations of their own devising, but they also had to report their work—an integral part of the scientific process. During the semester, students gave brief oral PowerPoint presentations to their lab sections summarizing their groups' progress, challenges, accomplishments, and next steps. In the early weeks, the presentations focused on refinement of the hypothesis being tested, as well as continuing incorporation of background information drawn from pertinent journal articles. Intermediate weeks often dealt with data analysis and statistics. The final week's oral presentations summarized results and conclusions. The ensuing discussions helped groups with formulation of the final research paper and scientific poster.

Every student was expected to participate in all reporting activities, which were consistent with what practicing scientists do professionally. Students were assessed on these deliverables for each half-semester project. Emphasis was placed on collaboration, reportage, and constructive criticism. Each student not only was required to give an oral presentation, but also was expected to participate in discussion of the presentations. Getting first-year students to criticize their peers is difficult, no matter how constructive the criticism. Consequently, the teaching team had to establish an ethos of "we're all in this together, and everyone can benefit from receiving and giving constructive criticism." In doing so, the undergraduate peer facilitator was especially helpful, because he or she was the least threatening and most relatable member of the teaching team. Each group of students was responsible for preparing a poster on their research project for display in a high-traffic area outside the laboratory classrooms at the end of the semester during a public scientific poster session held the week of final exams.

Students' lab grades were not dependent on the results of their research. Both oral and poster presentation activities were graded as "pass/fail." Avoiding letter grades minimized speakers' anxiety and individual competition among group members. For final, course letter grades, each student had to keep a laboratory notebook, write a one-page background summary of a relevant publication found through a literature search, and report their research in a standard (i.e., Introduction, Methods, Results, Discussion) scientific paper limited to two pages of text plus any needed additional pages of literature citations, tables, and figures with appropriate legends. A draft of the final research paper was thoroughly critiqued by the faculty instructor before a final revision was submitted for grading. Before the first paper draft was due, it was especially helpful to distribute a "model" paper conforming to the required format, written by the instructor. Other graded assessments, such as closed-book quizzes, take-home calculations, and problem sets, were included at the discretion of instructors. An example syllabus is available in Supplemental Table S2.

Laboratory Classrooms. The construction of two specially designed laboratory classrooms contributed to the success of UMARL. The rooms feature four custom-built, desk-height, octagonal worktables for student research groups, comfortable chairs on casters, full Internet connectivity, two laptop computers per table, a permanently mounted computer projector, and both standard and electronic whiteboards. The always-available computer projector facilitated the required student presentations, and at-hand laptops and the electronic whiteboard encouraged an emphasis on quantitative methods. The worktables helped to facilitate the collaborative group-work nature of the research projects, which would have been more difficult in traditional teaching lab classrooms with stools along the sides of stand-up-height, long, rectangular workbenches that impeded students facing and conversing with their group.

Course Enrollment

The majority of students who took UMARL were first-year undergraduates with majors in biology, chemistry, neuroscience, biochemistry, and other STEM disciplines. UMARL students' non–STEM majors included classics, business, political science, and psychology. Participant selection for these labs changed since the first implementation of the course in 2004. Originally, incoming freshmen majoring in biology or marine biology were rank-ordered by Scholastic Aptitude Test (SAT) or ACT scores, and every second student (effectively a stratified random group) was sent a letter of invitation to participate in UMARL. Because students had to have enough proficiency in mathematics to be able to do the graphical and statistical analyses required, only students with an SAT quantitative score at or above 550 or an ACT score of 26 or greater were invited. For the third year of implementation, we expanded the invitation list to include students on a premedical or preveterinary track, but not majoring in biology or marine biology. Responses to invitations were positive. For example, for the Fall 2007 semester, of 189 incoming freshmen who were sent invitations, 60% responded, with 111 (98.2% of respondents) accepting and only two declining. Before registration for the second semester, we offered students enrolled in the first-semester lab the opportunity to continue, and then opened enrollment as space allowed. Typically, students chose in which section to enroll based on their course schedule, and their majors were not considered in their assignments to lab sections and, therefore, research themes. Although the instructors for each section were known during enrollment, the research themes to be covered in a particular section where not explicitly stated in each semester's course listings. A substantial majority (137/213 = 64%) of first-semester invitees over the first 3 years (Fall 2004 through Spring 2007) of UMARL implementation chose to

continue for the second semester. Although we did not collect data on attrition, anecdotal evidence suggests the other 36% did not continue because they had scheduling conflicts or no requirement for a second semester of introductory biology lab in their program (such as the biomedical engineering major) or because some students did not like the uncertainty of real, authentic research.

In Fall 2008, UMARL was incorporated into a new program, the Advanced Program for Integrated Science and Mathematics (PRISM). The program's curriculum takes an integrative, interdisciplinary approach to teaching the basic sciences. PRISM students are block-scheduled and complete their introductory science coursework together during the first 2 years of undergraduate study. To qualify for PRISM, students are required to have a minimum SAT quantitative score of 720 (or ACT equivalent). Although most PRISM students are life science majors, that is not a requirement for admission to the program. PRISM students filled UMARL sections, with a few other students enrolled as space allowed.

Evaluation of UMARL

Student Learning Gains. To assess student engagement, we used the Classroom Undergraduate Research Experiences (CURE) survey (Lopatto, 2010) as an indirect measure of course learning gains. This survey is a helpful tool to evaluate the impact of CUREs on students, even though the survey is not formally validated (Hanauer and Dolan, 2014). It enables a comparison to other lab courses via national averages of CURE survey responses. At the end of each semester from Fall 2008 to Fall 2014 (inadvertently omitting Spring 2009), we invited UMARL students to participate in the CURE postcourse national survey administered online by David Lopatto and Leslie Jaworski from Grinnell College (www.grinnell.edu/academics/areas/psychology/assessments/cure-survey). The survey measures student self-perceived learning gains in course elements related to research skills, learning benefits, and attitudes toward science. Students rated their learning gains in course elements on a Likert scale (1 = no gain, through 5 = very large gain). Lopatto and Jaworski tabulated the mean response scores of our UMARL students and students nationally (national means included our students' responses), which they provided in a report each semester. Students in the comparison group participated in lab courses with different kinds of research experiences (i.e., structured inquiry, guided inquiry, open inquiry, or authentic research). The comparison group data were pooled from many institutions, the identities of which were not included in the reports provided by Lopatto and Jaworski.

Academic and Demographic Data for Summative Assessment. To assess objective outcomes of UMARL, we obtained from UM's Planning, Institutional Research, and Assessment Office the academic and demographic data for students who took traditional introductory biology labs (control group) and UMARL students (treatment group) from Fall 2004 through Spring 2014.

We requested the following data for each student: gender, race/ethnicity, SAT and/or ACT scores (including SAT-equivalent ACT scores), high school grade point average (GPA), term of matriculation, major at matriculation, introductory biology labs taken, terms when those labs were taken, participation in PRISM, research credits earned subsequent to introductory biology labs, term of graduation or last term enrolled, latest UM GPA, undergraduate degree earned, major at graduation or latest major, and graduation with honors.

After omitting students who might misrepresent either group (e.g., 1878 students who matriculated after Fall 2010 who did not have the opportunity to graduate in 4 years; Supplemental Table S3), our final sample included 3297 nontransfer, undergraduate degree-seeking students (2756 from traditional labs and 541 from UMARL), who matriculated between 2004 and 2010. Students who had taken at least one semester of UMARL were included with the treatment group, because they received the intervention (even if only for one semester).

Propensity Score Matching. We conducted propensity score matching to identify an unbiased, control group of students from UM's traditional introductory biology laboratory courses. These traditional labs were prescriptive, with students conducting different, stand-alone laboratory exercises each week. For each lab, students worked in groups of four to follow step-by-step instructions and answer questions given in the lab manual about their observations. The first-semester traditional biology lab covered topics such as the structure and division of cells (examination of slides under a light microscope), photosynthesis (extraction of chlorophyll from plants and analysis of light absorption properties), and developmental biology (observation of sea urchin embryo cleavage under a light microscope). The second semester of the traditional biology lab focused on biodiversity. In this lab, students examined slides under a light microscope, observed living and preserved specimens, and performed dissections to learn about various taxonomic groups.

Propensity scores correct for course selection bias and reflect the probability that a traditional lab student might have taken UMARL based upon a set of covariate descriptors of the actual UMARL participants (Rosenbaum and Rubin, 1983, 1984; West *et al.*, 2008). Propensity score matching resulted in a traditional lab control group with a similar mean SAT score to the UMARL group, accounting for the high SAT scores of PRISM students who took UMARL courses. We matched students using SAT score and STEM major at matriculation (0 = non-STEM, 1 = STEM) as covariates. We initially intended to use high school GPA as a covariate, but did not because of uncertain equivalencies and missing data. Gender, race/ethnicity, and graduation status (0 = not yet graduated, 1 = graduated) were used in exact matching. Gender and race/ethnicity are closely associated with student choice to major in STEM (Ma, 2011; Wang, 2013). Graduation status was used to ensure similar numbers of students in both groups for assessing outcomes associated with graduation. Using the Propensity Score Matching program for SPSS v. 3.0.4 (Thoemmes, 2011), we matched groups with the nearest-neighbor logistic regression as the algorithm with a caliper of 0.2 SDs of the logit of the estimated propensity score. The propensity score matching procedure resulted in traditional lab and UMARL groups of similar sizes, demographic makeup, SAT scores, numbers of students matriculated as STEM majors, and graduation status (Supplemental Table S4). The difference between propensity scores of both groups was reduced from 0.039 before matching to 0.002 after matching, a 96% overall reduction (Supplemental Figure S1).

Statistical Analyses

Student Learning Gains. We assessed student responses to all 21 learning gain items on the postcourse CURE survey taken by UMARL students and students nationally over 11 semesters (N = 280 and N = 38,160, respectively). We calculated Student's t tests using weighted means and standard deviations to compare differences in responses for all survey items related to course benefits. Differences were considered significant at α = 0.05.

Outcomes. We examined the effect of UMARL on outcome variables associated with student engagement in research, progress toward a STEM career, and academic achievement. We compared the propensity score matched groups using hierarchical logistic regressions, because our outcome data were binary (success or failure; students either enrolled in research or did not) and our observations of students fell into discrete categories, such as gender and ethnicity, in the traditional and UMARL groups. We then accounted for predictors that might influence results; for example, race is known to be associated with college GPA (Brown and Lee, 2005). Gender, race/ethnicity, and SAT score were known predictors included in the first step of the hierarchical models to account for effects correlated with outcomes. We checked for other potential predictors by examining correlations (Pearson's r) with each outcome variable before running hierarchical logistic regressions, and if significant, those predictors were added to the first step of the model. After all predictors were accounted for in step 1 (Supplemental Table S5), the hierarchical logistic model then investigated the likelihood of students who participate in UMARL having "success" (coded as "1") or not (coded as "0") compared with students in the traditional labs. This likelihood was provided as the odds ratio, or the ratio of the "odds" (i.e., the success vs. failure) of the UMARL group compared with that of the traditional lab group. Model fit was assessed using the Hosmer and Lemeshow goodness-of-fit test ($p > 0.05$, because the null hypothesis is that the chosen model is deemed appropriate) at each step of the regression (Peng *et al.*, 2002). The outcome variables we analyzed were whether students participated in subsequent individual research experiences (1) or not (0), graduated with a STEM major (1) or not (0), graduated in 4 years (1) or not (0), and graduated with honors (1) or not (0). We had planned to analyze undergraduate GPA at graduation (or latest GPA), but a generalized linear model was a poor fit for the data, so we did not proceed further.

Subsequent Individual Research Experiences. We used enrollment in independent research courses as a proxy for participation in subsequent individual research experiences. Although students also can obtain research experiences through lab internships, employment, and volunteering, that information is difficult to obtain. Research courses at UM allow students to earn academic credit for conducting independent research projects in faculty members' laboratories. We compared acquisition of research credits between the two groups of students by coding this variable (1 = with research credits, 0 = no research credits). In addition to gender, race/ethnicity, and SAT score, this logistic regression model included whether students matriculated as STEM majors in the first step of the regression (1 = STEM, 0 = non-STEM), because this variable was significantly correlated with whether a student acquired research credits or not (Pearson's $r = 0.143$, $p < 0.001$).

Major at Graduation. We looked at the likelihood that a student graduated with a STEM degree (1 = STEM, 0 = non-STEM) after participation in UMARL versus traditional biology labs. Whether a student was a STEM major at matriculation and the number of semesters (either one or two) a student participated in traditional introductory biology labs or UMARL were variables significantly correlated with STEM degree at graduation (STEM major at matriculation: Pearson's $r = 0.251$, $p \leq 0.0001$; number of semesters: Pearson's $r = 0.135$, $p \leq 0.0001$). Thus, we accounted for these predictors in the first step of the hierarchical logistic model. For this outcome, we considered only students who had graduated.

Time to Graduation. We examined the likelihood that UMARL students graduated in 4 years or fewer versus traditional lab students (1 = 4 years or fewer, 0 = more than 4 years). The number of semesters of introductory biology lab taken (either one or two) and number of research credits acquired were significantly correlated with graduation in 4 years or fewer (number of semesters: Pearson's $r = 0.123$, $p \leq 0.0001$; number of research credits: Pearson's $r = 0.188$, $p \leq 0.0001$) and thus were included in the first step of the logistic regression. All propensity matched students, regardless of graduation status, were included in this analysis.

Graduation with Honors. We investigated whether UMARL students were more likely than traditional lab students to receive honorable distinctions at graduation (1 = graduation with honors, 0 = regular graduation). Graduation with honors included students who graduated with general honors, departmental honors, and/or university honors (cum laude, summa cum laude, or magna cum laude). Students who graduated with general honors were required to complete honors courses and maintain a GPA of 3.5 or higher, and those with departmental honors, of which there were only three in our sample, also were required to have a minimum GPA of 3.5 and to conduct 6 credit hours or more of independent study, senior thesis, senior project, or advanced courses. University honors distinctions are based on GPA percentiles within each graduating class. In the regressions for this outcome, we used only students who had graduated. Not surprisingly, graduation with honors was significantly correlated with undergraduate GPA (Pearson's $r = 0.708$, $p \leq 0.0001$), so we consider this outcome an indicator of student achievement.

RESULTS

Student Learning Gains

Students who participated in UMARL reported significantly greater learning gains than students nationally for all 21 course benefit items (Figure 1 and Supplemental Table S6). The largest mean differences between UMARL students and students nationally were for elements related to communicating scientific research (effective oral presentation and science writing) followed by elements related to perseverance and self-efficacy (tolerance for obstacles, self-confidence, and readiness for research; Figure 1 and Supplemental Table S6). UMARL students reported their greatest learning gains overall, however, for items related to the scientific research process, such as learning lab techniques (mean rating$_{weighted}$ = 4.21 $\pm$ 0.21 SD$_{weighted}$), ability to analyze data (4.19 $\pm$ 0.18), and skill in interpretation of results (4.07 $\pm$ 0.19), as well as items related
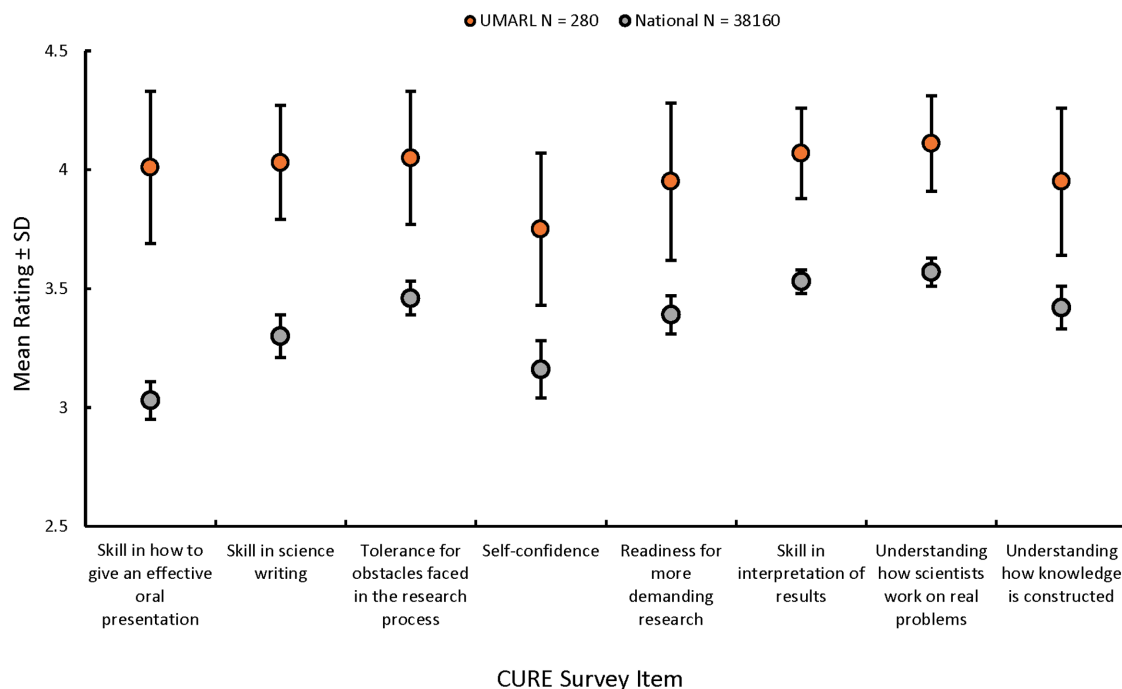
FIGURE 1. Weighted mean ratings (± weighted SD; 1 = no gain, 5 = very large gain) of self-reported benefits (all 21 surveyed course benefit gains are listed in Supplemental Table S6) from students in UMARL (*N* = 280) and students nationwide (*N* = 38,160). Items are arranged in diminishing order from largest difference in mean rating. Students were surveyed using the CURE survey by David Lopatto and Leslie Jaworski from Grinnell College. The national weighted means include the UMARL data.

to understanding the nature of science, such as understanding how scientists work on real problems (4.11 ± 0.2), understanding that scientific assertions require supporting evidence (4.08 ± 0.26), and understanding science (4.04 ± 0.27). Additionally, students indicated large learning gains in tolerance for obstacles (4.05 ± 0.28), skill in science writing (4.03 ± 0.24), and skill in how to give an effective oral presentation (4.01 ± 0.32). Although still significantly greater than students nationally, UMARL students reported moderate learning gains for items not directly addressed by the UMARL courses, including confidence in potential as a teacher (3.25 ± 0.29) and learning to work independently (3.43 ± 0.36).

Propensity Score Matching

Propensity score matching resulted in two similarly composed groups of students ($N_{traditional}$ = 532, N_{UMARL} = 541). Overall, the two groups were 61.6% female and 38.4% male with an ethnicity profile of 15.2% Asian or Pacific Islander, 0.5% American Indian or Alaska Native, 8.3% Black, 16.25% Hispanic, 52% white, 1.65% two or more races, and 6.15% unknown or missing (Supplemental Table S4). The mean SAT score was 1316 for both groups, with 92.25% of students having declared a STEM major at matriculation and 82.75% having graduated (Supplemental Table S4). The UMARL group was 28.5% PRISM students, with no PRISM students in the traditional labs.

Subsequent Individual Research Experiences

Participation in UMARL had a significant association with the likelihood of having subsequent individual research experi-

ences, as measured by enrollment in research course credits (β = 0.421, Wald χ^2 = 8.164; *p* = 0.004; Table 3 and Supplemental Table S5). After accounting for the significantly positive effects of STEM major at matriculation and SAT score, the odds of earning research course credit for UMARL students were 1.53 times greater than the odds for students from the traditional labs (Table 4). Overall, 38% of UMARL students and 29% of traditional lab students earned research credits (Figure 2A), and for both groups, the majority participated in one to three courses (Figure 2B).

Major at Graduation

UMARL was significantly associated with the likelihood that students graduated with a STEM major (β = 0.445, Wald χ^2 = 7.493; *p* = 0.006; Table 3 and Supplemental Table S5). The odds of UMARL students graduating with a STEM degree were 1.56 times greater than the odds for traditional lab students after accounting for ethnicity, SAT score, STEM major at matriculation, and the number of semesters of introductory biology lab taken (Table 4). Overall, 75.6% of UMARL and 69.2% of traditional lab students graduated with a STEM degree (Figure 3).

Of the variables included in step 1 of the logistic regression, STEM major at matriculation was the most highly correlated with graduation with a STEM degree. The odds of graduating with a STEM degree were 6.2 times greater for students who declared a STEM major at matriculation than the odds for students who declared a non-STEM major at matriculation (Table 4). Students who participated in two semesters of introductory biology lab had greater odds of graduating with a

TABLE 3. Hierarchical logistic regression model statistics at each step for models testing whether UMARL influenced 1) student engagement in research by acquiring research course credits, 2) graduating with a STEM degree, and student achievement reflected by 3) graduating in 4 or fewer years, and 4) graduating with honors

Outcome variable	Step	Predictor[a]	−2LL[b]	Pseudo-R^2	$\chi^{2\,c}$
Acquired research course credits	1	Control	1076.060	0.062	57.142
	2	Group	1067.836	0.098	65.366
Graduation with a STEM degree	1	Control	956.469	0.139	89.666
	2	Group	948.889	0.150	97.246
Graduation in four or fewer years	1	Control	1328.678	0.109	89.625
	2	Group	1317.644	0.122	100.658
Graduation with honors	1	Control	1013.412	0.215	151.939
	2	Group	1001.247	0.231	164.103

[a]Control step includes variables as described in the *Statistical Analyses* section.
[b]−2LL = −2*log likelihood; pseudo-R^2 = Nagelkerke R^2 estimate of effect size; χ^2 = chi-square value from omnibus tests of model coefficients.
[c]All reported χ^2 values in the table are significant at $p \le 0.0001$.

STEM degree than students who took only one semester (Table 4). SAT score also was significantly correlated with graduating with a STEM degree; high SAT scores were associated with STEM degree at graduation. Although both SAT score and STEM major at matriculation influenced graduating with a STEM major, they were not correlated with one another (Pearson's $r = 0.052$, $p = 0.088$). Students of Asian or Pacific Islander ethnicity had greater odds of graduating with a STEM major compared with the odds for students of other ethnicities, regardless of lab type (traditional or UMARL; Table 4). This was the only effect of ethnicity we found in our analyses.

Time to Graduation
UMARL was significantly associated with time to graduation (β = 0.447, Wald $\chi^2 = 10.924$; $p = 0.001$; Table 3 and Supplemental Table S4). The odds of graduating in 4 years or fewer for UMARL students were 1.56 times greater than the odds for traditional lab

students after accounting for the effects of gender, SAT score, research credit courses, and number of semesters of introductory biology lab (Table 4). In the UMARL group, 81% of students graduated in 4 years or fewer, as did 70.3% of traditional lab students (Figure 3). Regardless of lab type, females had greater odds than males to graduate in 4 years or fewer (Table 4). SAT score was positively correlated with this outcome (Table 4). Students with research course credits had greater odds of graduating in 4 years or fewer than students without research course credits, and those who took two semesters of introductory biology lab also had greater odds of graduating in 4 years or fewer than those who only took one semester of lab (Table 4).

Graduation with Honors
UMARL was positively associated with graduating with honors, even after accounting for SAT score (β = 0.535, Wald $\chi^2 = 12.011$; $p = 0.001$; Tables 3 and 4 and Supplemental Table S5).

TABLE 4. Significant predictors of outcomes ($p < 0.01$) and their respective odds ratios and 95% odds ratio confidence intervals (CI) from hierarchical logistic regressions[a]

Outcome	Step	Predictor	Odds ratio	Odds ratio 95% CI
Subsequent individual research experience	0	Constant	0.002	—
	1	SAT score	1.003	1.001–1.004
	1	STEM major at matriculation	8.424	2.995–23.692
	2	Participation in UMARL	1.534	1.141–2.034
STEM major at graduation	0	Constant	0.004	—
	1	Asian or Pacific Islander	2.947	1.402–6.196
	1	SAT score	1.002	1.001–1.004
	1	STEM major at matriculation	6.234	3.487–11.145
	1	Number of semesters in introductory biology lab	1.958	1.391–2.756
	2	Participation in UMARL	1.560	1.135–2.145
Graduation in 4 years or fewer	0	Constant	0.025	—
	1	Female	1.657	1.268–2.166
	1	SAT score	1.002	1.001–1.003
	1	Research course credit	1.487	1.282–1.723
	1	Number of semesters in introductory biology lab	1.731	1.308–2.291
	2	Participation in UMARL	1.564	1.200–2.040
Graduation with honors	0	Constant	0.000	—
	1	SAT score	1.008	1.006–1.009
	2	Participation in UMARL	1.707	1.262–2.311

[a]The effect of each predictor on the outcomes was accounted for in step 1 of the hierarchical logistic regression, before determining the effect of participating in UMARL (step 2) on the likelihood of gaining individual research experiences, graduating with a STEM major, graduating in 4 years, and graduating with honors.
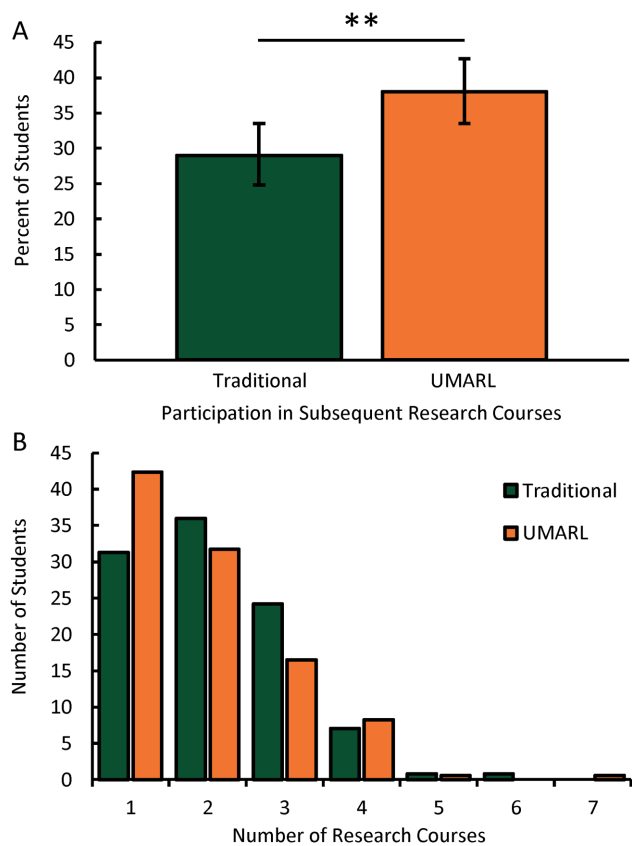
FIGURE 2. Percentage of students (± 95% binomial confidence interval) from UM traditional introductory biology laboratory (*N* = 441) or UMARL (*N* = 447) courses who engaged in subsequent individual research by acquiring research course credits (A) and the distribution of the number of research courses for each group (B). Asterisks indicate that participation in the UMARL labs significantly increased the odds ratio of students participating in subsequent research experiences (*p* < 0.01) after accounting for their major at matriculation and SAT scores in hierarchical logistic regressions.

The odds for UMARL students to graduate with honors were 1.71 times greater than the odds for students who took the traditional labs (Table 4). Because graduation with honors was significantly associated with GPA, we inferred that UMARL students also likely had greater odds than traditional lab students of graduating with a high GPA. Overall, 10% more UMARL students graduated with honors than students who took traditional labs (Figure 3).

DISCUSSION

Our hypothesis that UMARL would have a positive influence on students' learning gains compared with students nationally was supported by our results. Additionally, UMARL students had greater odds of participating in individual research, graduating with a STEM degree, graduating with honors, and graduating within 4 years than a matched group of UM's traditional biology lab students, as we initially predicted. Our approach of exposing first-year students to authentic research in laboratory classrooms may have helped students to understand scientific research, provided an experience in which students can be scientists, and may have fostered students' research and psychosocial skills.

Authenticity Is Important

Our CURE model embodies authentic research, which we propose is vital for the positive outcomes that we observed for UMARL students in comparison to UM's traditional biology lab students. Authenticity was created through students' freedom to ask their own, novel research questions, design their own experiments, and participate in discovery of knowledge new to science. These course attributes gave students responsibility for their own learning and a sense of personal investment that may have led to a high level of project ownership (Hanauer *et al.*, 2012; Hanauer and Dolan, 2014). Thus, students developed self-efficacy, including confidence in their ability to conduct research and a sense of belonging to the scientific community, which may be key factors for student persistence in STEM (Estrada *et al.*, 2011; Hernandez *et al.*, 2013; Robnett *et al.*, 2015).

UMARL required students to do what real scientists do, and so students' self-reported learning gains reflected their attainments in scientific research skills and, especially, their understanding of how scientists work. Our course design involved collaboration and iteration (with each switch of research theme), which were highly effective in teaching students to communicate science orally and in writing. Our CURE survey results suggest that UMARL students learned research skills while conducting research, much as graduate students do (Sadler *et al.*, 2010).

Improvements in student self-efficacy that we detected may be necessary for the positive benefits achieved by UREs (Sadler *et al.*, 2010; Adedokun *et al.*, 2013). UMARL students self-reported gains in excess of national means for both "tolerance for obstacles faced in the research process" and "self-confidence." When their data were inconclusive or unexpected, UMARL students discovered their ability to persevere. Moreover, repetition of activities with each new research theme, a fundamental tenet of our approach, also likely contributed to increased self-efficacy, because, as the saying goes "Practice makes perfect." Through four different research projects over the course of the academic year, students experienced the universal utility of the scientific research method and gained confidence in their ability to successfully conduct research.

Our analyses found significant benefits of UMARL, even with inclusion of students who took only one semester of the lab, despite the fact that some studies have found a correlation between the benefits of research experiences and time spent participating in research (Sadler *et al.*, 2010; Rodenbusch *et al.*, 2016). Regardless of their number of semesters of UMARL, approximately four UMARL students for every three traditional lab students had subsequent UREs, as has been determined qualitatively for other CUREs (Bascom-Slack *et al.*, 2012; Hanauer *et al.*, 2012; Harvey *et al.*, 2014; Kowalski *et al.*, 2016). Those studies reported that 75–92% of students had UREs, however, while only 38% of our UMARL students had them. Nevertheless, our analysis used research course credit as an objective measure of subsequent UREs, resulting in a highly conservative estimate.
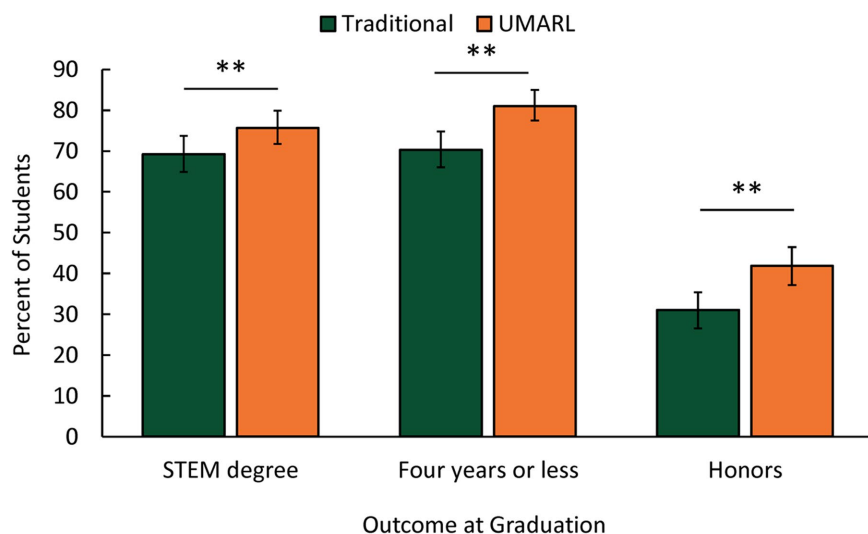
FIGURE 3. Percentage of students (± 95% binomial confidence interval) from UM traditional introductory biology laboratory or UMARL courses who obtained a STEM degree at graduation ($N_{\text{traditional}} = 441$; $N_{\text{UMARL}} = 447$), graduated in 4 years or fewer ($N_{\text{traditional}} = 532$; $N_{\text{UMARL}} = 541$), and graduated with an honors distinction ($N_{\text{traditional}} = 441$; $N_{\text{UMARL}} = 447$). Asterisks indicate that participation in UMARL significantly increased the odds ratio of students meeting each outcome at graduation ($p < 0.01$) after accounting for all other significant predictors in hierarchical logistic regressions.

Even though 92% of students in both of our propensity score matched groups declared STEM majors at matriculation, likely because of their precollege experiences (Tai *et al.*, 2006; Crisp *et al.*, 2009), UMARL still was positively associated with graduating with a STEM degree. Such an effect can be especially important in keeping minorities and females from leaving science (Nagda *et al.*, 1998; Hurtado *et al.*, 2009; Schultz *et al.*, 2011; Bangera and Brownell, 2014; National Academies of Sciences, Engineering, and Medicine, 2015).

Our results accord with findings of Rodenbusch *et al.* (2016) that participation in all three FRI courses had a significant effect on the probability of graduating with a STEM major and graduating within 6 years. Students who participated in only one or two FRI courses, however, were not any more likely than non-FRI students to have those outcomes. In contrast, regardless of the number of semesters for which students participated in UMARL, we found that UMARL students had increased odds of graduating in 4 years or fewer versus the odds for traditional lab students. Rodenbusch *et al.* (2016) found no effect of FRI courses on GPA at graduation, but our analysis did indicate a positive effect of UMARL on academic achievement, as reflected by graduation with honors.

UMARL has the hallmarks of authentic research described by Brownell and colleagues (2012): development of student-generated research questions with unknown answers, a focus on one or two research questions during the duration of the course, implementation of experimental designs that are developed as part of the course, collaboration, and presentation of research results. Similarly, SEA-PHAGES, the GEP, and FRI bring authenticity to course-based research via these elements, but through a different course structure and implementation model. For example, SEA-PHAGES focuses on one research theme with instructors who have been trained in phage discovery and bioinformatics

for comparative phage genome analyses, which may or may not be their areas of expertise. In contrast, FRI students conduct research under a single theme led by faculty upon whose research programs the themes are based, which is similar to UMARL. Nevertheless, UMARL students conduct research in four different research themes, whereas SEA-PHAGES, GEP, and FRI students typically conduct research under one research theme for an extended time. The differences among these CUREs illustrate that authenticity can be established through a variety of course models.

Implementation Considerations

Implementing any CURE is not without challenges. Lack of time for faculty to develop research projects is a major barrier to implementation of CUREs across different institution types (Kloser *et al.*, 2011; Lopatto *et al.*, 2014; Spell *et al.*, 2014; Kowalski *et al.*, 2016), but we circumvented this by basing course research on faculty instructors' own research programs and by giving students freedom to develop their own projects. Other CUREs have been developed around faculty research, but most focus on one research area (e.g., Bascom-Slack *et al.*, 2012; Kloser *et al.*, 2013; Harvey *et al.*, 2014; Olimpo *et al.*, 2016), whereas our model exposes students to multiple areas. The advantage is that application of the scientific process to different areas reinforces research skills and students' understanding of the process of science. Furthermore, if students do not connect with one research area, they may appreciate another.

Using faculty instructors' own research programs as bases for UMARL courses reduced our funding needs and kept the labs financially sustainable. In general, the cost of CUREs ranges widely; the few studies that have reported costs estimated from as little as $20 to as much as $500 per student (Rowland *et al.*, 2012; Burnette and Wessler, 2013; Harvey *et al.*, 2014; Russell *et al.*, 2015; Dolan, 2016). Each UMARL section had a supplies budget of $1000, about $42 per student. Although laboratory equipment requirements could make start-up costs steep, we avoided that by using faculty instructors' own specialized lab equipment. We purchased general equipment such as balances and micropipettes for the UMARL courses with grant and departmental funds, but the research areas of faculty were too diverse to purchase all the equipment necessary for each research theme. As to unanticipated supplies needed for particular student-designed projects, faculty often could supply small amounts of necessary reagents from their research lab, thereby avoiding the need to buy a large amount of a specialized reagent and wait for an order to arrive.

UMARL student access to sophisticated analytical instruments (e.g., microplate reader, CN analyzer, mass spectrometer, transmission electron microscope) that included trips to faculty members' labs, had an unexpected beneficial effect. It effectively emphasized to students that they were doing authentic research in a far more convincing way than just telling them

that. The use of sophisticated equipment also may contribute to publication-quality data (Weaver *et al.*, 2008), a component of authentic CUREs. In contrast, the use of simple, unsophisticated measuring equipment, such as student-constructed respirometers, which we tried in one of the very first labs in an attempt to demystify complex lab equipment, was abhorrent to students and characterized by them as worse than what they did in their high school labs. Being allowed to use (with close supervision) complicated, often computer-controlled, research lab instruments invariably excited and engaged students.

Because each teaching team oversees four different research projects simultaneously, teaching UMARL can be demanding. A graduate teaching assistant and undergraduate peer facilitator alleviated some of that intensity for the faculty instructors, but the three-person teaching teams did make the labs personnel intensive. Furthermore, faculty were given reduced teaching loads to help compensate for the time demands of teaching UMARL. Despite the challenges, faculty were attracted to teaching the labs, partly because they could effectively "prescreen" and then recruit students for individual research in their labs beginning in those students' second years. Moreover, research projects in UMARL often provided novel results that were informative for faculty research and that could be a basis for UREs. At least one such research project led to a publication with a UMARL student (e.g., Janos *et al.*, 2008), for which the research conducted as part of the course was continued and expanded. In UMARL, faculty instructors most comfortable with the unscripted nature of the labs engendered the most authentic research learning experiences. Faculty must be comfortable with free-form, impromptu instruction and able to "think on their feet" in order to guide students with interpretation of unexpected results and problem solving.

Student frustration and dissatisfaction with the authentic research labs could be high when things did not work as anticipated. Even though students' grades were not related to the conclusiveness of their experiments, first-year students fresh from high school cookbook labs are not prepared for setbacks. The teaching team had to help students understand that failures of technique, problem solving, and revision of experimental design are all realities of "real" science (and, while somewhat painful, can greatly impress upon students the authenticity of their experiences).

Another frustration for both students and faculty was the time constraint of 6 or 7 weeks (a half-semester for each research theme) of 3-hour labs. Not only were students trying to conclude a research project, but especially in the first semester, they also were developing proficiency with presentation software, mastering computational and spreadsheet skills, gaining an ability to interpret simple statistical tests, and learning the standard organization of a scientific paper and how to search the scientific literature. In response, some faculty and teaching assistants convened optional, "open" labs in lieu of office hours and facilitated by the availability of the UMARL classrooms for exclusive use by the UMARL course.

UMARL was designed knowing that there must be a compromise between the number of different research themes covered and the amount of time available to obtain conclusive data. Students often remarked that they wished a particular research project could run for an entire semester, but others, who did not like a particular theme, looked forward to the next one. Overall, our experience is that it is best to "leave them wanting more." Thereby, students learn that research seldom is conclusive. Moreover, exposure to multiple research areas provides students with breadth. Four different research projects in two semesters also involve repetition that reinforces research skills and builds students' self-efficacy.

Through feedback from students, we learned that many felt UMARL should confer more than one credit, because it is "more work" than the traditional biology labs. The traditional labs frequently have quizzes and homework assignments, so we suspect that the "more work" of the authentic research labs meant that they required student initiative, creativity, and analytical thinking. Because students might have had to work outside the assigned lab period (e.g., for off-campus sample collection or group poster preparation), additional credit could be justified. Nevertheless, we kept the authentic research labs on par with the traditional labs at 1 credit hour so as not to disrupt the requirements of the biology or other STEM majors.

Over our more than a decade of offering UMARL, we have found that the most successful and authentic research themes are relevant to topics of societal and scientific importance and have a high multiplication product of the numbers of independent and dependent variables that students can control and assess. High multiplication products allow very many possible investigations. Successful research themes have not relied on highly developed or extensive taxonomic, methodological, or conceptual expertise. They had easy access to adequate experimental subjects, rapid responses to experimentally manipulated variables, and closeness to the faculty member's research. In contrast, our least successful projects involved excessive student downtime (e.g., waiting for access to an expensive, single, critical instrument or piece of equipment that all student groups needed to use; waiting for a slow response to experimental treatments), too much complexity, or too many steps. Although we have found these research theme design considerations indispensable, we also have found many research faculty eager and able to meet them.

CONCLUSIONS

UMARL-style labs are neither difficult nor fiscally expensive to implement, and their benefits can shape students' academic paths and perhaps their eventual career choices, because CUREs are thought to have the potential to increase the number and diversity of STEM graduates entering STEM careers (NRC, 2003; American Association for the Advancement of Science, 2011; PCAST, 2012). The UMARL model, however, is inescapably personnel intensive. Nevertheless, the UMARL approach is relatively easily transportable to other STEM disciplines and institutions with available faculty researchers. Recently, we have extended these authentic research labs to introductory chemistry and to a nearby community college.

We suggest that the student learning gains and objective outcomes we have reported vouchsafe that our UMARL approach is well worth its costs in personnel, time, and effort. The weighted means of all 21 learning gains self-reported by UMARL students exceeded national means for every item, evincing strong student appreciation for our labs. Our outcome data are objective and embody relatively large (more than 500 students in each group), propensity score matched groups. Although the inclusion among our treatment group of the

highly qualified PRISM students (who would be expected to persist in STEM majors and to excel academically) could have influenced our observed outcomes, our traditional biology lab comparison group students were similarly qualified in academic background and potential (i.e., SAT score and matriculation with a STEM major). Taken together, those highly qualified groups might have minimized our chances of finding detectable, beneficial effects of UMARL. Despite that, the four outcomes we analyzed had significant odds ratios of 1.5–1.7, indicating strong merit of the UMARL model to contribute to student academic success and persistence in STEM.

ACKNOWLEDGMENTS

REFERENCES

Adedokun, O. A., Bessenbacher, A. B., Parker, L. C., Kirkham, L. L., & Burgess, W. D. (2013). Research skills and STEM undergraduate research students' aspirations for research careers: Mediating effects of research self-efficacy. *Journal of Research in Science Teaching*, *50*(8), 940–951. doi: 10.1002/tea.21102

Adedokun, O. A., Zhang, D., Parker, L. C., Bessenbacher, A., Childress, A., & Burgess, W. D. (2012). Understanding how undergraduate research experiences influence student aspirations for research careers and graduate education. *Journal of College Science Teaching*, *42*(1), 82–90.

American Association for the Advancement of Science. (2011). *Vision and change in undergraduate biology education: A call to action*. Washington, DC.

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. doi: 10.1187/cbe.14-01-0004

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. doi: 10.1187/cbe.14-06-0099

Bascom-Slack, C. A., Arnold, A. E., & Strobel, S. A. (2012). Student-directed discovery of the plant microbiome and its products. *Science*, *338*(6106), 485–486. doi: 10.1126/science.1215227

Blanchard, M. R., Southerland, S. A., Osborne, J. W., Sampson, V. D., Annetta, L. A., & Granger, E. M. (2010). Is inquiry possible in light of accountability? A quantitative comparison of the relative effectiveness of guided inquiry and verification laboratory instruction. *Science Education*, *94*(4), 577–616.

Brown, P. L., Abell, S. K., Demir, A., & Schmidt, F. (2006). College science teachers' views of classroom inquiry. *Science Education*, *90*(5), 784–802.

Brown, R. P., & Lee, M. N. (2005). Stigma consciousness and the race gap in college academic achievement. *Self and Identity*, *4*(2), 149–157, doi: 10.1080/13576500444000227

Brownell, S. E., Kloser, M. J., Fukami, T., & Shavelson, R. (2012). Undergraduate biology lab courses: Comparing the impact of traditionally based "cookbook" and authentic research-based courses on student lab experiences. *Journal of College Science Teaching*, *41*(4), 36–45.

Brownell, S. E., Kloser, M. J., Fukami, T., & Shavelson, R. (2013). Context matters: Volunteer bias, small sample size, and the value of comparison groups in the assessment of research-based undergraduate introductory biology lab courses. *Journal of Microbiology and Biology Education*, *14*(2), 176–182. doi: 10.1128/jmbe.v14i2.609

Buck, L. B., Bretz, S. L., & Towns, M. H. (2008). Characterizing the level of inquiry in the undergraduate laboratory. *Journal of College Science Teaching*, *38*(1), 52–58.

Burnette, J. M., & Wessler, S. R. (2013). Transposing from the laboratory to the classroom to generate authentic research experiences for undergraduates. *Genetics*, *193*, 367–375.

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. doi: 10.1187/cbe.14-10-0167

Crisp, G., Nora, A., & Taggart, A. (2009). Student characteristics, pre-college, college, and environmental factors as predictors of majoring in and earning a STEM degree: An analysis of students attending a Hispanic serving institution. *American Educational Research Journal*, *46*(4), 924–942. doi: 10.3102/0002831209349460

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. doi: 0.1152/advan.90112.2008

Dolan, E. (2016). *Course-based undergraduate research experiences: Current knowledge and future directions (Paper Commissioned for the Committee on Strengthening Research Experiences for Undergraduate STEM Students, Board on Life Sciences, Division of Earth and Life Studies)*. Washington, DC: National Academies of Sciences, Engineering, and Medicine.

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. doi: 10.3102/0002831213482038

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. doi: 10.1037/a0020743

Falchikov, N., & Boud, D. (1989). Student self-assessment in higher-education—a metaanalysis. *Review of Educational Research*, *59*(4), 395–430. doi: 10.3102/00346543059004395

Graham, M. J., Frederick, J., Byars-Winston, A., Hunter, A. B., & Handelsman, J. (2013). Increasing persistence of college students in STEM. *Science*, *341*(6153), 1455–1456. doi: 10.1126/science.1240487

Hanauer, D. I., & Dolan, E. L. (2014). The Project Ownership Survey: Measuring differences in scientific inquiry experiences. *CBE—Life Sciences Education*, *13*(1), 149–158. doi: 10.1187/cbe.13-06-0123

Hanauer, D. I., Frederick, J., Fotinakes, B., & Strobel, S. A. (2012). Linguistic analysis of project ownership for undergraduate research experiences. *CBE—Life Sciences Education*, *11*(4), 378–385. doi: 10.1187/cbe.12-04-0043

Hanauer, D. I., Graham, M. J., SEA-PHAGES, Betancur, L., Bobrownicki, A., Cresawn, S. G., ... & Hatfull, G. F. (2017). An inclusive research education community (iREC): Impact of the SEA-PHAGES program on research outcomes and student learning. *Proceedings of the National Academies of Science USA*, *114*(51), 13531–13536. doi: 10.1073/pnas.1718188115

Hansen, M., & Birol, G. (2014). Longitudinal study of student attitudes in a biology program. *CBE—Life Sciences Education*, *13*, 331–337.

Harrison, M., Dunbar, D., Ratmansky, L., Boyd, K., & Lopatto, D. (2011). Classroom-based science research at the introductory level: Changes in career choices and attitude. *CBE—Life Sciences Education*, *10*(3), 279–286.

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. doi: 10.1187/cbe.14-05-0089

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. doi: 10.1037/a0029691

Hunter, A. B., 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. doi: 10.1002/sce.20173

Hurtado, S., Cabrera, N. L., Lin, M. H., Arellano, L., & Espinosa, L. L. (2009). Diversifying science: Underrepresented student experiences in structured research programs. *Research in Higher Education*, 50(2), 189–214. doi: 10.1007/s11162-008-9114-7

Janos, D. P., Garamszegi, S., & Beltran, B. (2008). Glomalin extraction and measurement. *Soil Biology & Biochemistry*, 40(3), 728–739. doi: 10.1016/j.soilbio.2007.10.007

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 in phage discovery and genomics for first-year undergraduate students. *MBio*, 5(1), e01051-13. doi: 10.1128/mBio.01051-13

Kloser, M. J., Brownell, S. E., Chiariello, N. R., & Fukami, T. (2011). Integrating teaching and research in undergraduate biology laboratory education. *PLoS Biol*, 9(11), e1001174. doi: 10.1371/journal.pbio.1001174

Kloser, M. J., Brownell, S. E., Shavelson, R. J., & Fukami, T. (2013). Effects of a research-based ecology lab course: A study of nonvolunteer achievement, self-confidence, and perception of lab course purpose. *Journal of College Science Teaching*, 42(3), 72–81.

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. doi: 10.1187/cbe.16-02-0089

Linn, M. C., Palmer, E., Baranger, A., Gerard, E., & Stone, E. (2015). Undergraduate research experiences: Impacts and opportunities. *Science*, 347(6222). doi: 10.1126/science.1261757

Lopatto, D. (2004). Survey of Undergraduate Research Experiences (SURE): First findings. *Cell Biology Education*, 3(4), 270–277. doi: 10.1187/cbe.04-07-0045

Lopatto, D. (2007). Undergraduate research experiences support science career decisions and active learning. *CBE—Life Sciences Education*, 6(4), 297–306. doi: 10.1187/cbe.07-06-0039

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

Lopatto, D., Alvarez, C., Barnard, D., Chandrasekaran, C., Chung, H. M., Du, C., … & Elgin, S. C. (2008). Undergraduate research. Genomics Education Partnership. *Science*, 322(5902), 684–685. doi: 10.1126/science.1165351

Lopatto, D., Hauser, C., Jones, C. J., Paetkau, D., Chandrasekaran, V., Dunbar, D., … & Elgin, S. C. R. (2014). A central support system can facilitate implementation and sustainability of a classroom-based undergraduate research experience (CURE) in genomics. *CBE—Life Sciences Education*, 13(4), 711–723. doi: 10.1187/cbe.13-10-0200

Ma, Y. (2011). College major choice, occupational structure and demographic patterning by gender, race and nativity. *Social Science Journal*, 48(1), 112–129.

Nagda, B. A., Gregerman, S. R., Jonides, J., von Hippel, W., & Lerner, J. S. (1998). Undergraduate student–faculty research partnerships affect student retention. *Review of Higher Education*, 22(1), 55–72.

National Academies of Sciences, Engineering, and Medicine. (2015). *Integrating discovery-based research into the undergraduate curriculum: Report of a convocation*. Washington, DC: National Academies Press.

National Academies of Sciences, Engineering, and Medicine. (2017). *Undergraduate research experiences for STEM students: Successes, challenges, and opportunities*. Washington, DC: National Academies Press.

National Research Council (NRC). (2000). *Inquiry and the National Science Education Standards: A guide for teaching and learning*. Washington, DC: Center for Science, Mathematics, and Engineering Education.

NRC. (2003). *BIO2010: Transforming undergraduate education for future research biologists*. Washington, DC: National Academies Press.

Olimpo, J. T., Fisher, G. R., & DeChenne-Peters, S. E. (2016). Development and evaluation of the *Tigriopus* course-based undergraduate research experience: Impacts on students' content knowledge, attitudes, and motivation in a majors introductory biology course. *CBE—Life Sciences Education*, 15(4), ar72. doi: 10.1187/cbe.15-11-0228

Otero, V., Finkelstein, N., McCray, R., & Pollock, S. (2006). Who is responsible for preparing science teachers? *Science*, 313(5787), 445–446. doi: 10.1126/science.1129648

Peng, C. Y. J., So, T. S. H., Stage, F. K., & John, E. P. S. (2002). The use and interpretation of logistic regression in higher education journals: 1988–1999. *Research in Higher Education*, 43(3), 259–293.

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.

Robnett, R. D., Chemers, M. M., & Zurbriggen, E. L. (2015). Longitudinal associations among undergraduates' research experience, self-efficacy, and identity. *Journal of Research in Science Teaching*, 52(6), 847–867. doi: 10.1002/tea.21221

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), 10. doi: 10.1187/cbe.16-03-0117

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41–55. doi: 10.1093/biomet/70.1.41

Rosenbaum, P. R., & Rubin, D. B. (1984). Reducing bias in observational studies using subclassification on the propensity score. *Journal of the American Statistical Association*, 79(387), 516–524.

Rowland, S., Pedwell, R., Lawrie, G., Lovie-Toon, J., & Hung, Y. (2016). Do we need to design course-based undergraduate research experiences for authenticity? *CBE—Life Sciences Education*, 15(4), ar79. doi: 10.1187/cbe.16-02-0102

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, 46–62.

Russell, J. E., D'Costa, A. R., Runck, C., Barnes, D. W., Barrera, A. L., Hurst-Kennedy, J., … & Schlueter, M. (2015). Bridging the undergraduate curriculum using an integrated course-embedded undergraduate research experience (ICURE). *CBE—Life Sciences Education*, 14(1), ar4.

Russell, S. H., Hancock, M. P., & McCullough, J. (2007). Benefits of undergraduate research experiences. *Science*, 316(5824), 548–549. doi: 10.1126/science.1140384

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. doi: 10.1002/tea.20326

Schultz, P. W., Hernandez, P. R., Woodcock, A., Estrada, M., Chance, R. C., Aguilar, M., & Serpe, R. T. (2011). Patching the pipeline: Reducing educational disparities in the sciences through minority training programs. *Educational Evaluation and Policy Analysis*, 33(1), 95–114. doi: 10.3102/0162373710392371

Seymour, E., & Hewitt, N. M. (2000). *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. (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. doi: 10.1187/09-11-0087

Shaffer, C. D., Alvarez, C. J., Bednarski, A. E., Dunbar, D., Goodman, A. L., Reinke, C., … & Elgin, S. C. (2014). A course-based research experience: How benefits change with increased investment in instructional time. *CBE—Life Sciences Education*, 13(1), 111–130. doi: 10.1187/cbe-13-08-0152

Shortlidge, E. E., & Brownell, S. E. (2016). How to assess you CURE: A practical guide for instructors of course-based undergraduate research

experiences. *Journal of Microbiology & Biology Education*, *17*(3), 399–408. doi: https://doi.org/10.1128/jmbe.v17i3.1103

Spell, R. M., Guinan, J. A., Miller, K. R., & Beck, C. W. (2014). Redefining authentic research experiences in introductory biology laboratories and barriers to their implementation. *CBE—Life Sciences Education*, *13*(1), 102–110. doi: 10.1187/cbe.13-08-0169

Tai, R. H., Sadler, P. M., & Mintzes, J. J. (2006). Factors influencing college science success. *Journal of College Science Teaching*, *36*(1), 52–56.

Theobald, R., & Freeman, S. (2014). Is it the intervention or the students? Using linear regression to control for student characteristics in undergraduate STEM education research. *CBE—Life Sciences Education*, *13*(1), 41–48. doi: 10.1187/cbe-13-07-0136

Thiry, H., & Laursen, S. L. (2009). *Student outcomes from undergraduate research: An evaluation of three academic year and summer undergraduate research programs in the life sciences at the University of Colorado, Boulder, 2007–2008*. Retrieved August 27, 2017, from www.colorado.edu/eer/sites/default/files/attached-files/bsi_nih_ursurveyreport2009.pdf

Thoemmes, F. (2011). An SPSS R menu for propensity score matching. Retrieved January 4, 2017, from https://sourceforge.net/projects/psmspss/files/?

Wang, X. L. (2013). Why students choose STEM majors: Motivation, high school learning, and postsecondary context of support. *American Educational Research Journal*, *50*(5), 1081–1121. doi: 10.3102/0002831213488622

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. doi: 10.1187/cbe.11-03-0028

West, S. G., Duan, N., Pequegnat, W., Gaist, P., Des Jarlais, D. C., Holtgrave, D., … & Mullen, P. D. (2008). Alternatives to the randomized controlled trial. *American Journal of Public Health*, *98*(8), 1359–1366. doi: 10.2105/AJPH.2007.124446