Scholarship of Teaching and Learning in Psychology

Cultivating Inclusive Research Experiences Through Course-Based Curriculum

Viji Sathy, Chris L. Strauss, Mahfuz Nasiri, A. T. Panter, Kelly A. Hogan, and Bryant L. Hutson

Online First Publication, August 13, 2020. http://dx.doi.org/10.1037/stl0000215

Scholarship of Teaching and Learning in Psychology

ISSN: 2332-2101

http://dx.doi.org/10.1037/stl0000215

Cultivating Inclusive Research Experiences Through Course-Based Curriculum

Viji Sathy, Chris L. Strauss, Mahfuz Nasiri, A. T. Panter, Kelly A. Hogan, and Bryant L. Hutson
University of North Carolina at Chapel Hill

Course-based undergraduate research experiences (CURE) courses provide a systematic way for students to engage in a sustained, hypothesis-driven research experience within a classroom setting. While CURE courses are well defined in the natural sciences, our university expanded the implementation and assessment of CUREs to other disciplines, such as psychology and other social sciences (Auchincloss et al., 2014). We examine the generalizability of a commonly used CURE assessment, the 17-item Laboratory Course Assessment Survey (LCAS; Corwin, Runyon, Robinson, & Dolan, 2015) to disciplines outside of biology, including nonbiology natural sciences and social sciences, and to larger-scale classrooms. Multiple-indicator multiple-cause models (MIMIC) were used to evaluate measurement invariance and differential item functioning of the three LCAS factors (collaboration, discovery and relevance, iteration) by discipline and class size. Findings show that the LCAS items perform outside of the context of small biology laboratory courses and in large-scale psychology courses, such as introductory statistics and research methods. Additionally, to understand the factors that relate to CUREs within psychology, we identified effects of several demographic, psychosocial, academic factors on LCAS scores.

Keywords: CURE, undergraduate research, inclusive teaching, high-impact practices, differential item functioning

Research is essential to augment society's capacity to gather and utilize resources to address challenges and embrace new opportunities. Whether it involves uncovering new treatments for illnesses or advancing ideas in technology, the application of research techniques to reconstruct and renovate prior knowledge into new knowledge with broad relevance is an essential practice. To promote greater exposure to authentic, hands-on research experiences among undergraduate students, undergraduate research experiences have been introduced widely in higher education institutions (Olsen & Riordan, 2012).

While various approaches to providing undergraduate research experiences have been introduced and implemented successfully, these experiences generally took the form of mentored internships in faculty-led research groups or other apprenticeship models which limited the number of students who could participate (Kinkead, 2012; Wei & Woodin, 2011). Over time, in part because of these successes, there has been an increased emphasis on the importance of students pursuing research opportuni-

Viji Sathy, Department of Psychology and Neuroscience, Office of Undergraduate Education, University of North Carolina at Chapel Hill; Chris L. Strauss, Department of Psychology and Neuroscience, Institutional Research & Assessment, University of North Carolina at Chapel Hill; Mahfuz Nasiri, Institutional Research & Assessment, Gillings School of Global Public Health, University of North Carolina at Chapel Hill; A. T. Panter, Department of Psychology and Neuroscience, Office of Undergraduate Education, University of North Carolina at Chapel Hill; Kelly A. Hogan, Department of Biology, Office of Undergraduate Education, University of North Carolina at Chapel Hill; Bryant L. Hutson, Institutional Research & Assessment, University of North Carolina at Chapel Hill.

Correspondence concerning this article should be addressed to Viji Sathy, Department of Psychology and Neuroscience, Office of Undergraduate Education, University of North Carolina at Chapel Hill, CB 3270 Davie Hall, Chapel Hill, NC 27599-3270. E-mail: viji_sathy@unc.edu

ties (American Association for the Advancement of Science, 2011).

Course-based undergraduate research experiences (CUREs) courses offer more scalable and accessible research experiences to undergraduate students. CUREs are generally credit-bearing courses in a curriculum that allow students to simply enroll, as long as they meet prerequisites. While characteristics of CUREs were initially defined in the context of biology lab courses of 20 or fewer students (Auchincloss et al., 2014; Corwin, Runyon, Robinson, & Dolan, 2015), CUREs are adaptable to a multitude of disciplines, class enrollment sizes, and overall course structures.

Undergraduates who participate in CUREs demonstrate positive gains and outcomes similar to gains in mentored research experiences. Such gains include heightened interests in science and research (Harrison, Dunbar, Ratmansky, Boyd, & Lopatto, 2011; Ward, Clarke, & Horton, 2014), self confidence in the field of science (Brownell & Tanner, 2012), and higher conceptual understanding (Shaffer et al., 2010; Ward et al., 2014).

CUREs are defined by curricula that engage students in five key areas (Auchincloss et al., 2014; Mordacq, Drane, Swarat, & Lo, 2017):

1. Using scientific processes/scientific method. Students utilize the process of science in which includes formulating hypotheses, designing experiments, selecting methods, collecting and interpreting data, and communicating findings and limitations.
2. Discovery. Students ask questions and conduct research in which the answer(s) are unknown to students, faculty, and current field. This provides opportunities for students to learn where knowledge comes from and to impact the world by producing their own.
3. Collaboration. Students work collaboratively in any or all parts of the process of science, such as critiquing experimental designs or presenting findings to an audience. Authentic research is strengthened from diverse perspectives, and CUREs allow students to gain experience working in team settings to benefit from their peers' perspectives.
4. Iteration. Students revise work to address any limitations or issues that arise during the process of science, whether that be in narrowing down the scope of a novel question, or finding a better method of analysis. Through iteration, students learn to try, fail, and try again, ultimately equipping them with tools to approach a diverse array of scenarios.
5. Broad relevance. Students contribute novel perspectives and findings that may impact a discipline or a community. Students communicate the relevance through publications, reports, and/or presentations.

The Laboratory Course Assessment Survey (LCAS) is a 17-item survey instrument developed to differentiate CUREs from other laboratory courses by measuring students' perceptions of three design features of lab courses: (a) collaboration, (b) discovery and relevance, and (c) iteration (Corwin et al., 2015; see Appendix for LCAS items). In the initial Corwin, Graham, and Dolan (2015) study, the LCAS differentiated CUREs from traditional labs in discovery and iteration and provided a scale to compare CURE characteristics across biological lab courses.

At UNC-Chapel Hill, we have built a program to expand CURE courses across the sciences, social sciences, and arts and humanities. Faculty developing these courses engage in a year-long faculty learning community to develop courses around the five characteristics of CURE courses including alignment to assessment objectives such as the LCAS. Although the LCAS was developed within a biological discipline across small lab sections, we have utilized the instrument to explore how CURE features vary across disciplines and course sizes. Our current study focuses on the effectiveness of CURE implementation in large enrollment psychology courses (Statistical Methods in Psychology and Research Methods in Psychology). We also aim to understand the generalizability of the LCAS assessment tool to CUREs beyond biology.

Method

The site of this study is a leading research university that has identified providing meaningful research experiences to undergraduates as a significant institutional priority. As part of a campus-wide initiative, the CURE model was introduced to a number of disciplines, including psychology. Faculty who created CURE courses participated in a year-long faculty learning community designed to incorporate principles of CUREs and foster alignment to program goals. Sathy, Nasiri, Strauss, and Hutson (2020) outline pedagogical approaches used to incorporate CUREs at scale in psychology.

Data were collected from a sample of n = 2,110 undergraduate students enrolled in 21 distinct CURE classes across multiple disciplines (biology, chemistry, physics, astronomy, geological sciences, psychology, geography, and political science). Surveys containing the LCAS items were distributed to CURE instructors at the end of each semester via Qualtrics and instructors provided survey links to their students.

The LCAS is comprised of three scales. The collaboration scale consists of six items developed to measure student perceptions of how often they are encouraged to work together, and provide and respond to feedback. This scale also addresses student experiences with reflecting upon and developing metacognition toward their work within their class. The scale incorporates four response options that address frequency of activities: 1 = *never*, 2 = *one or two times*, 3 = *monthly*, or 4 = *weekly*.

The discovery and relevance scale involves five items designed to capture the degree to which students perceive themselves as having opportunities in class to create novel knowledge in the discipline and provide support for findings based on data. The discovery/relevance scale incorporates a 6-point Likert scale (1 = *strongly disagree*, 2 = *disagree*, 3 = *somewhat disagree*, 4 = *somewhat agree*, 5 = *agree*, 6 = *strongly agree*).

Finally, the iteration scale addresses the degree to which students perceive themselves as having opportunities to revise or repeat their work to address problems or new questions that arise in their research, and to improve the validity of their own and others' results. Consisting of six items, the subscale incorporates a 6-point Likert scale (1 = *strongly disagree*, 2 = *disagree*, 3 = *somewhat disagree*, 4 = *somewhat agree*, 5 = *agree*, 6 = *strongly agree*).

Responses on items across the three scales are aggregated to create overall scores for collaboration, discovery and relevance, and iteration. A total LCAS score is comprised of the three combined scores. The possible range of scores for collaboration is 6–24, for discovery and relevance is 5–30, and for iteration is 6–36. The overall LCAS score may range from 17 to 90.

We analyzed the data to evaluate: (a) differential item functioning (DIF) across large versus small classes specifically within psychology CUREs; and (b) the impact of demographic factors on differences in LCAS scores in the psychology CUREs. All DIF analyses were conducted in Mplus using the WLSMV estimator given the ordinal response scale of LCAS items. The multivariate regression was conducted using Proc REG in SAS.

We evaluated DIF within a multiple indicator multiple cause modeling framework (MIMIC) using χ^2 difference tests to compare the model assuming no differentially functioning items to a model in which a direct path from the grouping variable to one of the items is included (Muthén, 1985, 1988, 1989). Statistical significance of the χ^2 statistic indicates item under consideration is noninvariant with regards to the grouping variable. This method is appropriate when group sizes are relatively small, and allows for use of estimators that can be used with binary or polytomous response scales, such as robust weighted least squares (Brown, 2015; Muthén, Kao, & Burstein, 1991). Further, it is analogous to IRT-LR-DIF using all other items as anchors (Thissen, Steinberg, & Wainer, 1988, 1993).

Results

Study 1. Evaluating the LCAS Within Psychology and Examining DIF

Across all CURE courses offered at the institution, confirmatory factor analysis indicated good fit for the three factors of LCAS and all items yielded statistically significant loadings. To evaluate LCAS within psychology, we took a subset of the sample students in CURE psychology courses all taught by the same instructor, to avoid differences dependent on instructor level characteristics. This sample consisted of all students from the original sample, enrolled in PSYC 210, PSYC 270, and PSYC 530 from the Fall 2017 to Spring 2019 semester (see Table 1). Within this sample, we were specifically interested in examining DIF across large versus small classes, given the LCAS was originally validated only in small laboratory cours-

Table 1
Psychology CURE Courses Included in Sample

Courses	Semesters taught	Total enrollment	# of respondents	Response rate
PSYC 210: Introduction to Statistics for the Behavioral Sciences	Spring 2018	174	120	68.97%
	Fall 2018	88	64	72.73%
	Spring 2019	220	185	84.09%
PSYC 270: Research Methods in Psychology	Spring 2017	231	185	80.08%
	Fall 2017	39	35	89.74%
	Spring 2018	40	28	70.00%
PSYC 530: Design and Interpretation of Psychological Research	Fall 2017	25	23	92.00%
	Fall 2018	31	29	93.55%

Note. For more information about these courses, see Sathy, Nasiri, Strauss, and Hutson (2020).

es. We were also interested in better understanding invariance of the measure across sex and race/ethnicity.

Overall, the model exhibited good fit, suggesting the original LCAS factors are well-defined in the sample of psychology students, $\chi^2(113) = 483.889$, CFI = .98, TLI = .98, RMSEA = .070, 90% CI [.064, .077]. All items loaded significantly on their respective factors and modification indices did not suggest the notable presence of cross-loadings.

We also evaluated DIF across large versus small psychology courses, under-represented minority (URM) students and non-under-represented minority (non-URM) students, male and female students, and first-generation students (FGS) and nonfirst generation students (non-FGS). Due to missing data, the final sample size for all subsequent DIF analyses involved 661 students. Items were evaluated individually for DIF by each grouping category using nested likelihood ratio tests. Results are presented in Table 2.

Focus on class size. All items flagged for noninvariance by class size in the likelihood ratio tests were then included in a subsequent model to determine the direction and significance of the potential source of DIF. The model fit well, $\chi^2(125) = 571.301$, $p < .05$, CFI = .98, TLI = .97, RMSEA = .073, 90% CI [.067, .080]. Results are presented in Table 3.

Focus on URM. Next, items flagged for noninvariance by URM status were examined by fitting a model including paths from URM status directly to these items. The model fit well, $\chi^2(125) = 509.006$, $p < .05$, CFI = .98, TLI = .98, RMSEA = .068, 90% CI [.062, .074]. Results are presented in Table 4. Most survey items were noninvariant across race/ethnicity suggesting this tool is effective in evaluating CURE experiences for students of different race/ethnicity backgrounds.

Focus on sex. All items were invariant across sex, suggesting no inherent sex differences in the capacity of the LCAS to assess CUREs when comparing male and female students, thus additional analyses in regards to sex were unnecessary.

Focus on First-Generation College Student Status. Finally, paths from first-generation status to all items flagged for DIF were included in a model to assess the direction and significance of potential noninvariance. The model fit well, $\chi^2(124) = 518.545$, $p < .05$, CFI = .98, TLI = .98, RMSEA = .069, 90% CI [.063, .076]. Results are presented in Table 5.

Study 2. Exploring Relationships of Academic and Psychosocial Factors to the LCAS

After evaluating the generalizability of the LCAS over a variety of groups, we examined how demographic, academic, and psychosocial factors impacted LCAS scores of Psychology students. LCAS scores were computed by summing all items in each factor as suggested by previous research. Any student who failed to respond to one of the items was dropped from the sample in order to assure reliable scores.

Preliminary evidence suggested many of the variables we initially tested had little impact on LCAS scores based on statistical significance and effect size, so we selected a final multivariate regression model using sex, URM status, first-generation status, year in school, and the

Table 2
Likelihood Ratio χ2 Difference Statistics for Detecting Differential Item Functioning (N = 661)

Item	Class size (82.8% large)		Race/ethnicity (24.7% URM)		Sex (82.5% female)		First-generation status (26.8% FGS)	
	χ^2 difference statistic	*p*-value	χ^2 difference statistic	*p*-value	χ^2 difference statistic	*p*-value	χ^2 difference statistic	*p*-value
Collaboration								
Collab 1	.47	.49	1.48	.22	3.72	.05	5.51	<.05
Collab 2	7.76	<.01	.54	.46	.04	.85	2.84	.09
Collab 3	.62	.43	.14	.71	.59	.44	.47	.49
Collab 4	5.32	<.05	6.72	<.05	2.45	.12	2.99	.08
Collab 5	3.15	.08	.33	.57	.30	.58	.30	.58
Collab 6	.36	.55	.03	.86	1.39	.24	1.75	.19
Discovery/ Relevance (DR)								
DR 1	2.50	.11	.22	.64	1.45	.23	6.80	<.01
DR 2	2.28	.13	4.14	<.05	2.21	.14	1.15	.28
DR 3	6.03	<.05	.64	.42	.35	.55	1.17	.28
DR 4	.11	.74	3.65	.06	.10	.75	2.96	.09
DR 5	.11	.75	2.44	.12	.65	.42	.96	.33
Iteration								
Iteration 1	2.05	.15	.22	.64	.37	.54	1.74	.19
Iteration 2	.97	.32	.00	1.00	.17	.68	1.01	.32
Iteration 3	21.06	<.001	1.56	.21	1.75	.39	.46	.50
Iteration 4	1.44	.23	.49	.49	.00	.96	2.69	.10
Iteration 5	4.64	<.05	.06	.43	.06	.80	4.29	<.05
Iteration 6	2.47	.12	.10	.75	.29	.59	.75	.39

Note. URM = under-represented minority; Collab = collaboration. All 210 courses were considered large courses as enrollment surpassed 60 students per course. PSYC 270 taught in the Spring 2017 semester was considered a large class while the other sections were coded as small. All sections of PSYC 530 were coded as small courses.

Project Ownership-Content (Hanauer & Dolan, 2014) scale as predictors of total LCAS score, collaboration, discovery/relevance, and iteration.

The Project Ownership-Content scale measures personal, community, and disciplinary aspects of ownership that students may feel toward their research work (Hanauer & Dolan, 2014). This scale has been shown to be sensitive to the differentiation between course-based research experiences and traditional courses (Hanauer, Graham, & Hatfull, 2016). While the LCAS measures students' perceptions of course attributes, the Project Ownership-Content scale focuses on students sense of engagement and agency toward the research project. As described by Hanauer and Dolan (2014), project ownership involves identifying connections be-

Table 3
Parameter Estimates for Analysis Including Items Flagged for DIF by Class Size

Item	Estimate	Standard error	EST/*SE*	*p*-value
C2. I was encouraged to reflect on what I was learning	.03	.13	.23	.822
C4. I was encouraged to help other students collect or analyze data	−.37	.14	−2.59	<.05
DR3. I was expected to formulate my own research questions or hypothesis to guide an investigation	−.42	.12	−3.53	<.001
I3. I had time to share and compare data with other students	−.65	.11	−5.83	<.001
I5. I had time to revise or repeat analyses based on feedback	−.20	.11	−1.91	.06

Note. C = collaboration; D/R = discovery/relevance; I = iteration. *SE* = standard error; DIF = differential item functioning.

Table 4
Parameter Estimates for Analysis Including Items Flagged for DIF by Race/Ethnicity

Item	Estimate	Standard error	EST/*SE*	*p*-value
C4. I was encouraged to help other students collect or analyze data	.20	.08	2.59	<.05
DR2. I was expected to conduct an investigation to find something previously unknown to myself, other students, and the instructor	−.13	.07	−2.04	<.05

Note. C = collaboration; DR = discovery/relevance; I = iteration. *SE* = standard error; DIF = differential item functioning.

tween students' previous experiences and current research efforts, developing agency and engaging others in pursuing research work; overcoming challenges in their research, and expressing excitement and positive emotions when achieving research goals. LCAS scores have been explored as a predictor of project ownership in a study by Corwin et al. (2018) involving undergraduates participating Freshman Research Initiative. In this study, project ownership was used as a predictor of LCAS score along with student attributes.

The Project Ownership-Content score is computed by summing the 10 ownership items and dividing by 10. Any student who failed to respond to any of the Project Ownership-Content items or LCAS items were dropped from the sample in order to assure sum scores were not deflated by missing data. Further, students with missing values for demographics were also dropped from the analysis, leading to a final sample size of $N = 610$ psychology students. Descriptive statistics for all variables are included in Table 6.

Multivariate regression was selected due to strong overall correlations among the dependent variables. The LCAS total score and all LCAS factors were all highly and significantly correlated with each other. Coefficients are listed in Table 7. Parameter estimates are listed in Table 8.

Results suggested a significant overall test predicting LCAS total scores, $F(7, 602) = 42.10$, $p < .0001$, $R^2 = 0.329$, adjusted $R^2 = 0.321$. The Project Ownership-Content score significantly predicted LCAS total scores above and beyond the effects of all other covariates, ($\beta = 8.686$, $p < .0001$). Further, both sophomores ($\beta = -5.724$, $p < .01$) and juniors ($\beta = -5.199$, $p < .05$) showed significantly lower total LCAS scores than first-year students. For all of the subscales (collaboration, discovery/relevance, and iteration) we see a similar pattern with Project Ownership-Content, for example, highly predictive, followed by some minor differences in year in school. For instance, seniors reported lower collaboration scores than first years, and sophomores and juniors reported lower iteration than first-year students. Given the particularly low effect size, results should be interpreted with caution.

Discussion

Our research suggests that the LCAS is a robust measure for undergraduate research ex-

Table 5
Parameter Estimates for Analysis Including Items Flagged for DIF by First-Generation

Item	Estimate	Standard error	Estimate/*SE*	*p*-value
C1. I was encouraged to discuss elements of my investigation with classmates or instructors	.16	.10	2.35	<.05
DR1. I was expected to generate novel results that were unknown to the instructor and that could be of interest to the broader scientific community or others outside of class	.16	.06	2.61	<.01
I5. I had time to revise drafts of papers or presentations about my investigation based on feedback	−.12	.06	−2.07	<.05

Note. SE = standard error; DIF = differential item functioning.

Table 6
Descriptive Statistics for LCAS Scores

Variable	*M*	*SD*	Min	Max
LCAS total	72.32	10.56	17	90
LCAS collaboration	19.42	4.02	6	24
LCAS discovery/relevance	24.68	4.20	5	30
LCAS iteration	28.23	5.37	6	36
PITS ownership	3.63	0.67	1	5

Note. LCAS = Laboratory Course Assessment Survey.

periences in course-based curriculum. Not only does the factor structure pattern hold similarly across nonbiology courses as they do for biology courses (as it was initially designed for), but it also demonstrates very little differential item functioning across class size and the student characteristics examined in this paper. Finally, we demonstrate that the LCAS correlates highly and positively with a measure of project ownership.

In our training of faculty who teach these psychology courses and other courses in our university-wide CURE program, we utilize the five characteristics of CUREs (use of scientific practices, discovery, broad relevance, collaboration, and iteration described by Auchincloss et al. (2014) as our framework for the definition of research). Yet rather than describing the first characteristic as "use of scientific practices" we say "use of evidence-based practices" to be more discipline-inclusive. Cohort after cohort, the faculty learning communities always agree that these five descriptors and the LCAS scales make sense across their own disciplines just by altering just a few words or giving concrete examples to each other (i.e., selecting archival texts in humanities vs. survey data in social sciences). Thus, whether we are talking about producing knowledge on chemical synthesis or how students learn, we find research can be defined by these common attributes and thus measured similarly across disciplines.

Where we did see minor differences may be due to implementation of CURE in nonlaboratory settings as well as class size. For instance, formulating one's own research hypothesis and having time to iterate a research question can be compromised in a large-scale implementation of a CURE. What is promising is that the majority of items show very little DIF for class size, indicating that CUREs may be scaled to large class sizes with minimal impact on students perceived research skills. Often core psychology courses such as introductory statistics and research methods are taught in a large-enrollment format, so it is assuring to note that with skilled instructors and support for development of CUREs, it's possible for nearly all undergraduate psychology majors to engage in research as part of required courses in the curriculum.

Furthermore, the research on impact on subgroups is also reassuring. In nearly all cases, there was no DIF across gender, ethnic underrepresented minority status, and first-generation status. In fact, when there were differences they were nearly all in a positive direction. For instance, for students with the same level of perceived collaboration, URM students were more likely to report higher frequency of collaboration when responding to the item "I was encouraged to help other students collect or analyze data," compared with non-URM students. In the case of first-generation college students, the three items that exhibited DIF: "I was encouraged to discuss elements of my investigation with classmates or instructors," "I was expected to generate novel results that were unknown to the instructor and that could be of interest to the broader scientific community or others outside of class," and "I had time to revise drafts of papers or presentations about my investigation based on feedback" were all higher for students who identified as first-generation college students.

Lastly, we found some fairly small, yet statistically significant effects predicting LCAS by year in school. Specifically, these were in the collaboration and iteration subscale. It's likely that the extent to which collaboration has oc-

Table 7
Correlation Coefficients for LCAS Total Score and Factors

	LCAS total	Collaboration	Iteration	Discovery/ Relevance
Collaboration	.66*			
Iteration	.86*	.34*		
Discovery/ Relevance	.79*	.28*	.56*	

Note. LCAS = Laboratory Course Assessment Survey (Corwin, Graham, & Dolan, 2015).
* $p < .001$.

Table 8
Parameter Estimates for Multivariate Regression Model Predicting LCAS Scores

Predictors	β	*SE*	95% CI	
Sex (82% female)				
LCAS total	1.07	2.918	−.74	2.88
Collaboration	.41	.41	−.39	1.23
Discovery/Relevance	.58	.38	−.17	1.32
Iteration	.07	.48	−.86	1.01
URM status (25% URM)				
LCAS total	−.99	.83	−2.61	.63
Collaboration	−.56	.37	−1.29	.17
Discovery/Relevance	−.85*	.34	−1.52	−.18
Iteration	.07	.48	−.43	1.26
First-generation status (27% FGS)				
LCAS total	−.19	.80	−1.76	1.38
Collaboration	−.28	.36	−.98	.43
Discovery/Relevance	.53	.33	−.14	1.16
Iteration	−.44	.41	−1.25	.38
Sophomore (22%)				
LCAS total	−5.72**	2.09	−9.83	−1.63
Collaboration	−1.45	.94	−3.29	.39
Discovery/Relevance	−.89*	.86	−2.58	.82
Iteration	−3.39**	1.08	−5.52	−1.27
Junior (40%)				
LCAS total	−5.20*	2.03	−9.18	−1.22
Collaboration	−1.68	.91	−3.47	.01
Discovery/Relevance	−.68	.84	−2.32	.96
Iteration	−2.83**	1.05	−4.89	−.78
Senior (35%)				
LCAS total	−3.69	2.04	−7.69	.31
Collaboration	−2.12*	.91	−3.91	−.33
Discovery/Relevance	.45	.84	−1.20	2.09
Iteration	−2.01	1.05	−4.07	.06
Ownership				
LCAS total	8.69***	.54	7.64	9.74
Collaboration	1.41***	.24	.94	1.88
Discovery/Relevance	3.06***	.22	2.63	3.50
Iteration	4.21***	.28	3.67	4.75

Note. LCAS = Laboratory Course Assessment Survey (Corwin, Runyun, et al., 2015).
* $p < .05$. ** $p < .01$. *** $p < .001$.

curred at the time a student is a senior may be why this is rated lower than for non-Seniors. Finally, regarding differences in iteration, we may see lower scores for sophomores and juniors as they are disproportionately enrolled in the introductory statistics course where iteration is least emphasized in the implementation of the CURE in favor of collaboration and discovery combined with course content.

Taken together these findings indicate how useful the LCAS is in assessing collaboration, iteration, and discovery in the context of a research-based course. Our findings expand upon earlier research and adds that the measure is psychometrically sound outside laboratory biology courses and translates very well to courses in psychology. Furthermore, the lack of DIF in class size and for certain demographic subgroups, suggests that implementing CUREs in psychology allows for broadening participation of research experiences by all students who enroll in these classes, some of which are core, required classes in psychology. By allowing students to engage in research through course-based enrollment, we are minimizing the barriers to research opportunity access as well as student's know-how to seek out valuable and enriching research opportunities. There are many students for whom this is very beneficial, such as those who discover psychology later in their undergraduate career, those who are unfamiliar with navigating research opportunities, those who may not realize the importance of conducting research both for understanding how we approach psychological science, and those who may discover their own interest in research, to name a few. Consider this quote from a student:

> "It was so inspiring to see a course like this one being offered. The availability of your class to everyone is such an important component, and fills a great need that the university was missing. It opens so many doors for students trying to explore research and having trouble getting into a lab or something because of lack of prior experience. I remember how hard it was for me and I thought it was really cool that there are people like you who recognized the obstacles and did something about it. I was reflecting on my semester earlier and I had not even realized how much I had learned in your class because I am so preprogrammed to think about memorized facts now when asked about what I learned. Though there were some facts I learned in this class, I mostly learned skills, which I am realizing more and more, are more important to learn all things considered . . . thank you for giving me the confidence to further explore and present research! I am doing research next semester as well where I will have to present it, but waiting until the last minute (right before graduation) stresses me out. I really hope more professors follow your lead on student inclusivity and student-centered classes, as this was one of the most different but most enjoyable classes I have had the opportunity to take here. Thank you again for designing and offering such a fulfilling and rewarding course!"

By incorporating CURES in psychology and using tools such as the LCAS to assess and track implementation, we are cultivating an inclusive research experience through curriculum. Those of us who teach in psychology or develop psychology curriculum can consider this an opportunity to introduce the high-impact practice of research to all the students in our major.

References

American Association for the Advancement of Science. (2011). *Vision and change: A call to action* [Final Report]. Washington, DC: American Association for the Advancement of Science.

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,* 29–40. http://dx.doi.org/10.1187/cbe.14-01-0004

Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). New York, NY: Guilford Press.

Brownell, S. E., & Tanner, K. D. (2012). Barriers to faculty pedagogical change: Lack of training, time, incentives, and . . . tensions with professional identity? *CBE Life Sciences Education, 11,* 339–346. http://dx.doi.org/10.1187/cbe.12-09-0163

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*. Advance online publication. http://dx.doi.org/10.1187/cbe.14-10-0167

Corwin, L. A., Runyon, C. R., Ghanem, E., Sandy, M., Clark, G., Palmer, G. C., . . . Dolan, E. L. (2018). Effects of discovery, iteration, and collaboration in laboratory courses on undergraduates' research career intentions fully mediated by student ownership. *CBE Life Sciences Education, 17,* 2. http://dx.doi.org/10.1187/cbe.17-07-0141

Corwin, L. A., Runyon, C., Robinson, A., & Dolan, E. L. (2015). The Laboratory Course Assessment Survey: A tool to measure three dimensions of research-course design. *CBE Life Sciences Education*. Advance online publication. http://dx.doi.org/10.1187/cbe.15-03-0073

Hanauer, D. I., & Dolan, E. L. (2014). The project ownership survey: Measuring differences in scientific inquiry experiences. *CBE Life Sciences Education, 13,* 149–158. http://dx.doi.org/10.1187/cbe.13-06-0123

Hanauer, D. I., Graham, M. J., & Hatfull, G. F. (2016). A Measure of College Student Persistence in the Sciences (PITS). *CBE Life Sciences Education*. Advance online publication. http://dx.doi.org/10.1187/cbe.15-09-0185

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,* 279–286. http://dx.doi.org/10.1187/cbe.10-12-0151

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

Mordacq, J. J., Drane, D. L., Swarat, S. L., & Lo, S. S. (2017). Development of course-based undergraduate research experiences Using a design-based approach. *Journal of College Science Teaching, 46,* 64–75. http://dx.doi.org/10.2505/4/jcst17_046_04_64

Muthén, B. O. (1985). A method for studying the homogeneity of test items with respect to other relevant variables. *Journal of Educational Statistics, 10,* 121–132. http://dx.doi.org/10.3102/10769986010002121

Muthén, B. O. (1988). Some uses of structural equation modeling in validity studies: Extending IRT to external variables using SIMS results. In H. Wainer & H. Braun (Eds.), *Test validity* (pp. 213–238). Hillsdale, NJ: Erlbaum.

Muthén, B. O. (1989). Latent variable modeling in heterogeneous populations. *Psychometrika, 54,* 557–585. http://dx.doi.org/10.1007/BF02296397

Muthén, B. O., Kao, C., & Burstein, L. (1991). Instructionally sensitive psychometrics: Application of a new IRT-based detection technique to mathematics achievement test items. *Journal of Educational Measurement, 28,* 1–22. http://dx.doi.org/10.1111/j.1745-3984.1991.tb00340.x

Olsen, S., & Riordan, D. G. (2012). *Engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics* [Report to the President]. Washington, DC: Executive Office of the President, President's Council of Advisors on Science and Technology.

Sathy, V., Nasiri, M., Strauss, C., & Hutson, B. (2020). The CURE for Broadening Participation in Undergraduate Researche. In T. M. Ober, E. Che, J. E. Brodsky, C. Raffaele, & P. J. Brooks (Eds.), *How we teach now: The GSTA guide to transformative teaching* (pp. 17–29). Retrieved from http://teachpsych.org/ebooks/howweteachnow-transformative

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,* 55–69. http://dx.doi.org/10.1187/09-11-0087

Thissen, D., Steinberg, L., & Wainer, H. (1988). Use of item response theory in the study of group difference in trace lines. In H. Wainer & H. Braun (Eds.), *Test validity* (pp. 147–169). Hillsdale, NJ: Erlbaum.

Thissen, D., Steinberg, L., & Wainer, H. (1993). Detection of differential item functioning using the parameters of item response models. In P. W. Holland & H. Wainer (Eds.), *Differential item functioning* (pp. 67–111). Hillsdale, NJ: Erlbaum.

Ward, J. R., Clarke, H. D., & Horton, J. L. (2014). Effects of a research-infused botanical curriculum on undergraduates' content knowledge, STEM competencies, and attitudes toward plant sciences. *CBE Life Sciences Education, 13,* 387–396. http://dx.doi.org/10.1187/cbe.13-12-0231

Wei, C. A., & Woodin, T. (2011). Undergraduate research experiences in biology: Alternatives to the apprenticeship model. *CBE–Life Sciences Education, 10,* 123–131. http://dx.doi.org/10.1187/cbe.11-03-0028

Appendix

Items From the Laboratory Course Assessment Survey (LCAS; Corwin, Runyun, et al., 2015)

Items		Scale type
Collaboration		
Please read the following statements and indicate how often they applied . . .		4-point frequency scale (*never, one or two times, monthly, weekly*)
C1	I was encouraged to discuss elements of my investigation with classmates or instructors	
C2	I was encouraged to reflect on what I was learning	
C3	I was encouraged to contribute my ideas and suggestions during class discussions	
C4	I was encouraged to help other students collect or analyze data	
C5	I was encouraged to provide constructive criticism to classmates and challenge each other's interpretation	
C6	I was encouraged to share the problems I encountered during my investigation and seek input on how to	

(*Appendix continues*)

Appendix (*continued*)

Items	Scale type
Discovery/Relevance	
Please indicate the extent to which you agree or disagree with each statement	6-point Likert agreement scale
DR1 I was expected to generate novel results that were unknown to the instructor and that could be of interest to the broader scientific community or others outside of class	
DR2 I was expected to conduct an investigation to find something previously unknown to myself, other students, and the instructor	
DR3 I was expected to formulate my own research questions or hypothesis to guide an investigation	
DR4 I was expected to develop new arguments based on data	
DR5 I was expected to explain how my work has resulted in new knowledge	
Iteration	
Please indicate the extent to which you agree or disagree with each statement	6-point Likert agreement scale
I1 I was expected to revise or repeat work to account for errors or fix problems	
I2 I had time to change the methods of the investigation if it was not unfolding as predicted	
I3 I had time to share and compare data with other students	
I4 I had time to collect and analyze additional data to address new questions or further test hypotheses that arose during the investigation	
I5 I had time to revise or repeat analyses based on feedback	
I6 I had time to revise drafts of papers or presentations about my investigation based on feedback	

Received January 16, 2020
Revision received June 11, 2020
Accepted June 15, 2020 ■