Supplemental Material

CBE—Life Sciences Education

Casper and Laporte

Supplemental Materials

A CURE Lab in Introductory Biology at a Regional Comprehensive University Negatively Impacts Student Success in the Associated Lecture Course Among Underrepresented Groups in Science

Anne M. Casper, Marianne Laporte

Representative example of traditional lab schedule for Introductory Biology: Cells & Molecules. This lab meets once a week, for four hours. The first hour is for a recitation, and the remaining three hours are for a hands-on lab activity.

Week	Date	Recitation topic	Lab Activity
1	Sep 9 - 15	Scientific method	Scientific method – walking, running, and heart rate Plant Fast Plants
2	Sep 16 - 22	(none)	Searching for primary literature Plant oats and clovers
3	Sep 23 - 29	Macromolecules	Osmosis and diffusion – dialysis bags
4	Sep 30 – Oct 6	Tips for success	Enzymes – potato catalase
5	Oct 7 - 13	Cellular respiration	Photosynthesis – leaf punches and bicarbonate
6	Oct 14 - 20	Cell division	Mitosis and meiosis – pop beads
7	Oct 21 - 27	(none)	Lab Practical #1
8	Oct 28 – Nov 3	DNA replication	Molecular Biology part 1 – restriction enzyme digest of plasmids
9	Nov 4 – Nov 10	Transcription, Translation	Molecular Biology part 2 – gel electrophoresis of digested plasmids
10	Nov 11 – Nov 17	Mendelian genetics	Mendelian genetics – phenotypes of crosses between Fast Plants
11	Nov 18 – Nov 24	No lab - Thanksgiving	No lab - Thanksgiving
12	Nov 25 – Dec 1	Natural selection	Natural selection – bacteria on plates with and without antibiotics Harvest oats and clover for drying
13	Dec 2 – Dec 8	Organismal interactions	Organismal interaction – evaluate growth of oats and clover potted in various combinations
14	Dec 9 - 15	(none)	Scientific Poster presentations of organismal interaction data
15	Dec 16 - 22		Lab Practical #2

Representative example of CURE schedule for Introductory Biology: Cells & Molecules. This lab meets twice a week, for two hours each time.

Week	Date	Discussion topic	Lab Activity
1a	Sep 9	Antibiotic crisis, Lab safety, Soil	Practice putting on/taking off PPE
1b	Sep 11	Sterile technique, Growth media, Lab notebook	Plate for bacterial colonies from instructor-provided soil sample
2a	Sep 16	Serial dilution	Learn micropipetting and serial dilution
2b	Sep 18	More serial dilution	Plate for colonies from student-provided soil sample Determine soil type (clay, sand, silt)
3a	Sep 23	Metabolic pathways, Auxotrophs	View plates from student soil Repeat best dilution for screening experiment
3b	Sep 25	Intro to scientific literature Structure of Introductions in primary literature	Count CFU/g on control plate Make Art Palette plate
4a	Sep 30	Microscopes	Practice using a microscope (prepared slides and pond water)
4b	Oct 2	Gram staining, Peptidoglycan	Gram staining with control bacteria
5a	Oct 7	Structure of Mat & Meth in primary literature	Identify antibiotic-producing microbes Streak to purify antibiotic-producing microbes
5b	Oct 9	(none)	Gram stain antibiotic-producing microbes Repeat streaks to purify microbes as needed
6a	Oct 14	Penicillin, ESKAPE pathogens	Screen antibiotic-producing microbes against panel of ESKAPE relatives
6b	Oct 16	(none)	Use microscopes to view Gram stains Repeat Gram staining as needed Create Agar art
7a	Oct 21	Structure of Results section in primary literature	View plates of screen against ESKAPE relatives Re-streak microbes that kill ESKAPE relatives
7b	Oct 23	Jeopardy! (for review)	Agar Art competition
8a	Oct 28	Lab Practical #1	Lab Practical #1
8b	Oct 30	Sterilization by autoclave Tour CMBB lab	Make media Re-streak to further purify antibiotic-producing microbes
9a	Nov 4	Structure of Discussion in primary literature Finding primary lit	Re-streak to further purify
9b	Nov 6	Intro to the TE database PCR	Enter info in online TE database Make frozen stocks of antibiotic-producing microbes PCR on antibiotic-producing microbes
10a	Nov 11	Abstracts and Titles in primary literature Gel electrophoresis	Gel electrophoresis Repeat PCR as needed
10b	Nov 13	ExoSap	Repeat gel electrophoresis as needed ExoSAP treatment of successful PCR samples
11a	Nov 18	Sanger sequencing	Prepare ExoSAP-treated samples for Sanger sequencing
11b	Nov 20	Peer reviewing process	In-class peer review of two other lab reports
12a	Nov 25	Analyzing chromatograms, BLAST	BLAST analysis of sequencing results In-class literature search to find articles related to BLAST results
12b	Nov 27	No class - Thanksgiving	No class - Thanksgiving
13a	Dec 2	What is a scientific poster	Add more info to online TE database In-class work time to create individual poster
13b	Dec 4	How to present a scientific poster	In-class work time to create group poster Peer-review of another table's individual posters,
14a	Dec 9	(none)	In-class work time to create group poster
14b	Dec 11	Jeopardy! (for review)	Group poster presentation
15a	Final Exam	Lab Practical #2	Lab Practical #2

Student demographics are similar in the traditional lab (Trad) and CURE in both experiments.

	Experiment #1 (Volunteer CURE)[b]	**Experiment #2 (Non-volunteer CURE)**
Semesters of survey data and lecture exam data collection [a]	Fall 2018 Winter 2019 Fall 2019 Winter 2020	Fall 2016 (traditional only) [c] Winter 2017 (traditional only) [c] Fall 2017 (traditional only) Winter 2018 (traditional only) Fall 2021 (CURE only)

		Winter 2022 (CURE only) [d] Fall 2022 (CURE only) [d]
Total Enrollment	Trad n=684 CURE n=239	Trad n=1091 CURE n=619
Median concurrent GPA (1st-3rd quartile)	Trad 3.220 (2.333 – 3.737) CURE 3.291 (2.338 – 3.795)	Trad 3.100 (2.300 – 3.700) CURE 3.190 (2.300 – 3.809)
Female (% of total enrollment)	Trad n=453 (66%) CURE n=167 (70%)	Trad n=724 (66%) CURE n=434 (70%)
BIPOC (% of total enrollment)	Trad n=208 (30%) CURE n=76 (32%)	Trad n=323 (30%) CURE n=218 (35%)
Pell Eligible (% of total enrollment)	Trad n=311 (45%) CURE n=115 (48%)	Trad n=487 (45%) CURE n=257 (42%)
First Generation (% of total enrollment)	Trad n=182 (27%) CURE n=62 (26%)	Trad n=295 (27%) CURE n=139 (22%)

[a] No data was collected during academic year 2020 – 2021 due to the impact of the COVID pandemic.
[b] Multiple sections of the CURE and the traditional lab were offered every semester of Experiment 1
[c] Lecture exam data was not collected from Fall 2016 or Winter 2017.
[d] Survey data was not collected from Winter 2022 2016 or Fall 2022.

Post-course survey. This survey consisted of all items from the Laboratory Course Assessment Survey (LCAS) and all items from the Project Ownership Survey (POS).

Survey instrument	**Sub-scale**	**Questions**	**Likert Scale Response options**
Laboratory Course Assessment Survey (LCAS)	Collaboration	In this course…. I was encouraged to discuss element of my investigation with	Never; One or two times; Monthly; Weekly

(Corwin et al. 2015)		classmates of instructors I was encouraged to reflect on what I was learning I was encouraged to contribute my ideas and suggestions during class discussions I was encouraged to help other students collect or analyze data I was encouraged to provide constructive criticism to classmates and challenge each other's interpretations I was encouraged to share the problems I encountered during my investigation and seek input on how to address them	
	Iteration	In this course…. I was expected to revise or repeat work to account for errors or fix problems I had time to change the methods of the investigation if it was not unfolding as predicted I had time to share and compare data with other students I had time to collect and analyze additional data to address new questions or further test hypotheses that arose during the investigation I had time to revise or repeat analyses based on feedback I had time to revise drafts of papers or presentations about my investigation based on feedback	Strongly disagree; Disagree; Somewhat disagree; Somewhat agree; Agree; Strongly agree
	Discovery/Relevance	In this course…. I was expected to generate novel results that are unknown to the instructor and that could be of interest to the broader scientific community or others outside of class I was expected to conduct an investigation to find something previously unknown to myself, other students, and the instructor I was expected to formulate my own research questions or hypotheses to guide an investigation I was expected to develop new arguments based on data I was expected to explain how my work has resulted in new scientific knowledge	Strongly disagree; Disagree; Somewhat disagree; Somewhat agree; Agree; Strongly agree
Project Ownership Survey (POS) (Hanauer and Dolan 2014)	Cognitive Ownership	My research will help to solve a problem in the world. My findings were important to the scientific community. I faced challenges that I managed to overcome in completing my research project. I was responsible for the outcomes of my research. The findings of my research project gave me a sense of personal achievement. I had a personal reason for choosing the research project I worked on. The research question I worked on was important to me. In conducting my research project, I actively sought advice and	Strongly disagree; Disagree; Neither disagree nor agree; Agree; Strongly agree

		assistance. My research project was interesting. My research project was exciting.	
	Emotional Ownership	[a] To what extent does the word delighted describe your experience of the laboratory course? [a] To what extent does the word happy describe your experience of the laboratory course? [a] To what extent does the word joyful describe your experience of the laboratory course? [b] To what extent does the word astonished describe your experience of the laboratory course? [b] To what extent does the word surprised describe your experience of the laboratory course? [b] To what extent does the word amazed describe your experience of the laboratory course?	Very slightly; Slightly; Moderate; Considerably; Very strongly

[a] If the Project Ownership Survey items load onto three factors as suggested in (Corwin et al. 2018), these items load on the "Enjoyment" sub-scale.

[b] If the Project Ownership Survey items load onto three factors as suggested in (Corwin et al. 2018), these items load on the "Surprise" sub-scale.

Power analysis of data sets from Experiment #1 (recruitment bias present) and Experiment #2 (no recruitment bias) indicates that most of the effect sizes for student sub-groups are of medium magnitude. Cohen's d (effect size) is shown by student subgroup for each experiment for the latent constructs of Collaboration, Iteration, and Discovery/Relevance on the Laboratory Course Assessment Survey (LCAS), the latent constructs of enjoyment and surprise on the Project Ownership Survey (POS), and for Lecture Exam score. Effect sizes that have a medium or large magnitude are shown in bold green font. All other effect sizes are small. Classification of the magnitude of effect size follows the classification for educational interventions proposed by Kraft (2020): less than 0.05 is small, 0.05 to less than 0.20 is medium, and 0.20 and above is large.

Student sub-group	Collaboration	Iteration	Discovery/ Relevance	Enjoyment	Surprise	Lecture Exam Score

Exp 1, Pell eligible only	**0.13**	**0.08**	0.00	**0.06**	**0.09**	0.03
Exp 1, BIPOC only	**0.07**	**0.07**	**0.11**	**0.18**	**0.19**	**0.08**
Exp 1, Both Pell eligible and BIPOC	0.04	0.02	0.02	**0.05**	**0.07**	**0.28**
Exp 2, Pell eligible only	0.02	**0.15**	0.01	0.03	0.03	0.04
Exp 2, BIPOC only	0.03	**0.08**	0.03	0.01	0.01	**0.12**
Exp 2, Both Pell eligible and BIPOC	**0.13**	0.00	**0.05**	**0.07**	**0.06**	**0.24**

Regression Model Selection for Survey Data Analysis

Model selection was carried out as recommended for discipline-based education research (E. Theobald, 2018). Since students are nested in lab sections, and lab sections are nested in semesters, we evaluated whether these would be appropriate to include as random effects in a multilevel regression model. We first calculated the intraclass correlation (ICC) for lab section and semester in both Experiment #1 and Experiment #2 for each student perception of each subscale on the LCAS (collaboration, iteration, or discovery/relevance) or the POS (enjoyment or surprise). All ICC values were <0.001.

Random Effect	ICC for collaboration	ICC for iteration	ICC for discovery relevance	ICC for enjoyment	ICC for surprise
Exp1: Lab Section	<0.001	<0.001	<0.001	<0.001	<0.001

Exp1: Semester	<0.001	<0.001	<0.001	<0.001	<0.001
Exp2: Lab Section	<0.001	<0.001	<0.001	<0.001	<0.001
Exp2: Semester	<0.001	<0.001	<0.001	<0.001	<0.001

The fixed effects that we tested included Lab Type (Traditional or CURE), student subgroup (Pell only, BIPOC only, Both, or Neither), and Sex (male or female). We built a fixed effects-only model that tests the hypothesis that lab type affects student perceptions, and that student subgroups are differentially affected by lab type (Model 1). We then fit additional models with all the possible combinations of random effects (Models 2 – 4):

Model 1: Student perception ~ Lab Type * Subgroup + Sex
Model 2: Student perception ~ Lab type * Subgroup + Sex + (1|Lab Section)
Model 3: Student perception ~ Lab type * Subgroup + Sex + (1|Semester)
Model 4: Student perception ~ Lab type * Subgroup + Sex + (1|Lab Section) + (1|Semester)

Comparing Models 1 – 4, Model 1 has the lowest AIC in all cases. Therefore, neither lab section nor semester was retained as a random effect. This is consistent with the low ICC values for these effects. We then tested an additional model, removing the fixed effect of sex since this is not explicitly part of the hypothesis we are testing.

Model 5: Student perception ~ Lab Type * Subgroup

The AIC value is reported below for each model. The lowest AIC value (that is, the best-fitting model) is shown in blue bold font; any models within ΔAIC = 2 of the lowest AIC value are shown in black bold.

Model	Collaboration Exp 1 AIC	Collaboration Exp2 AIC	Iteration Exp 1 AIC	Iteration Exp2 AIC	Discovery Relevance Exp 1 AIC	Discovery Relevance Exp2 AIC	Enjoyment Exp 1 AIC	Enjoyment Exp2 AIC	Surprise Exp 1 AIC	Surprise Exp2 AIC
Mod1	**322.9**	**234.7**	**410.9**	**341.1**	**421.8**	**370.2**	**371.7**	333.5	**378.8**	**320.0**
Mod2	332.0	241.4	417.4	345.5	426.1	372.9	381.0	339.8	388.1	327.4
Mod3	336.4	249.1	417.7	346.8	427.3	372.9	381.6	339.8	388.1	327.3
Mod4	334.0	243.4	419.4	347.5	427.8	374.5	383.0	341.8	390.1	329.3
Mod5	**321.6**	**236.2**	**410.8**	**339.3**	**421.0**	**369.1**	374.6	**332.1**	**377.2**	**318.1**

In both Experiment #1 and #2, Model 5 is the best fit in nearly all cases because it has the lowest AIC and is the simplest model. For consistency we have chosen to use Model 5 for all survey data for both experiments.

Regression Model Selection for Lecture Exams Data Analysis

The proportion of lecture exam points earned from the total possible was calculated for each student. To avoid having the exam score dependent variable trapped between 0 and 1, we took the natural log of the odds of exam score. The ln(odds of exam score) was used as the outcome variable in linear regression. Model selection was carried out as recommended for discipline-based education research (E. Theobald, 2018). We evaluated whether lab section, semester, lecture instructor, or lecture instructor by semester (since lecture instructors taught in multiple semesters) would be appropriate to include as random effects in a multilevel regression model. We calculated the intraclass correlation (ICC) for each random effect in both Experiment #1 and Experiment #2.

Random Effect	ICC for Exams, Experiment #1	ICC for Exams, Experiment #2
Lab Section	0.032	0.025

Semester	0.002	0.046
Lecture Instructor by Semester	<0.001	0.089

The fixed effects that we tested included Lab Type (Traditional or CURE), student subgroup (Pell only, BIPOC only, Both, or Neither), sex (male or female), and concurrent GPA. To reduce negative skew in the concurrent GPA, we reflected each value by subtracting from five, and then took the natural log of the reflected value. We built a fixed effects-only model that tests the hypothesis that lab type affects lecture exam score, and that student subgroups are differentially affected by lab type (Model 1). We then fit additional models with all the possible combinations of random effects (Models 2 – 8).

	Exp 1 AIC	Exp 2 AIC
Model 1: exams ~ Lab Type * Subgroup + Sex + GPA	**939**	**852**
Model 2: exams ~ Lab Type * Subgroup + Sex + GPA + (1|Lab Section)	967	877
Model 3: exams ~ Lab Type * Subgroup + Sex + GPA + (1|Semester)	968	873
Model 4: exams ~ Lab Type * Subgroup + Sex + GPA + (1|Lec Instruc by Semester)	968	862
Model 5: exams ~ Lab Type * Subgroup + Sex + GPA + (1|Lab Section) + (1|Semester)	969	871
Model 6: exams ~ Lab Type * Subgroup + Sex + GPA + (1|Lab Section) + (1| Lec Instruc by Semester)	969	860
Model 7: exams ~ Lab Type * Subgroup + Sex + GPA + (1|Semester) + (1| Lec Instruc by Semester)	970	864
Model 8: exams ~ Lab Type * Subgroup + Sex + GPA + (1|Lab Section) + (1|Semester) + (1| Lec Instruc by Semester)	971	862

For both Experiments #1 and #2, Model 1 has the lowest AIC. Therefore no random effects were retained in the bet-fitted model for Experiment #1 or #2, which is consistent with the low ICC values for these effects. Next, we fit additional models using the process of backwards model selection, removing one fixed effect at a time in order, starting with the fixed effect that has the smallest effect. If AIC increased upon removal, we retained that effect. These models and the AIC values for Experiments #1 and #2 are listed below.

	Exp 1 AIC	Exp 2 AIC
Model 1: exams ~ Lab Type * Subgroup + Sex + GPA	**939**	**852**
Model 9: exams ~ Lab Type * Subgroup + GPA	**940**	857
Model 10: exams ~ Lab Type + Sex + GPA	1052	919
Model 11: exams ~ Lab Type * Subgroup + Sex	1177	1053

For Experiment #2, Model 1 is the best fit because it has the lowest AIC. For Experiment #1, Model 1 and Model 9 have AIC values with equivalent fit. For consistency we have chosen to use Model 1 for both experiments.

Results from individual items on the Discovery/Relevance sub-scale of the LCAS

Horizontal bars indicate median and boxes represent the interquartile range for each item. The items included on this subscale are: In this course….

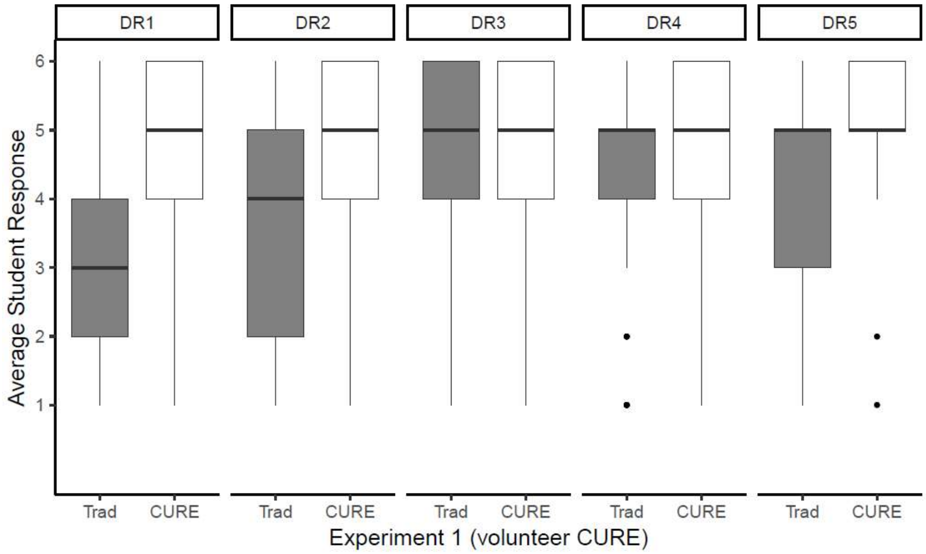
(DR1) I was expected to generate novel results that are unknown to the instructor and that could be of interest to the broader scientific community or others outside of class

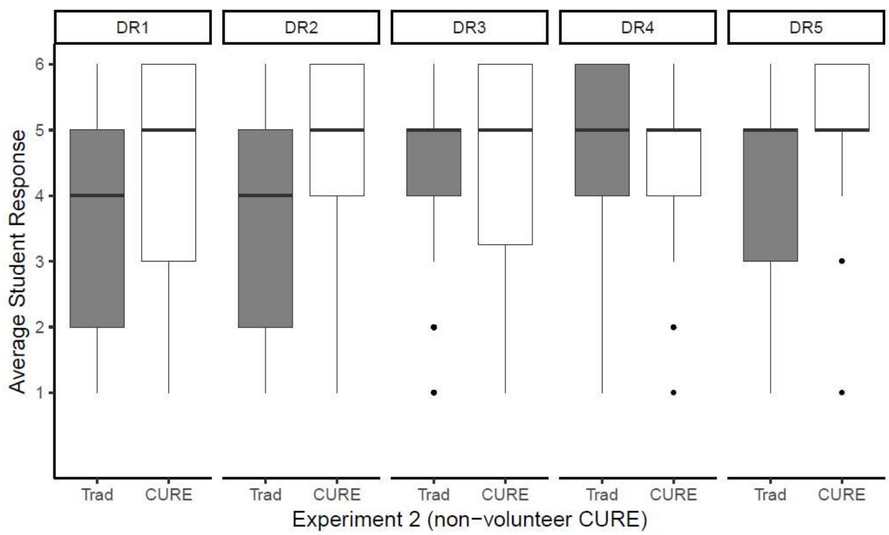
(DR2) I was expected to conduct an investigation to find something previously unknown to myself, other students, and the instructor

(DR3) I was expected to formulate my own research questions or hypotheses 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 scientific knowledge



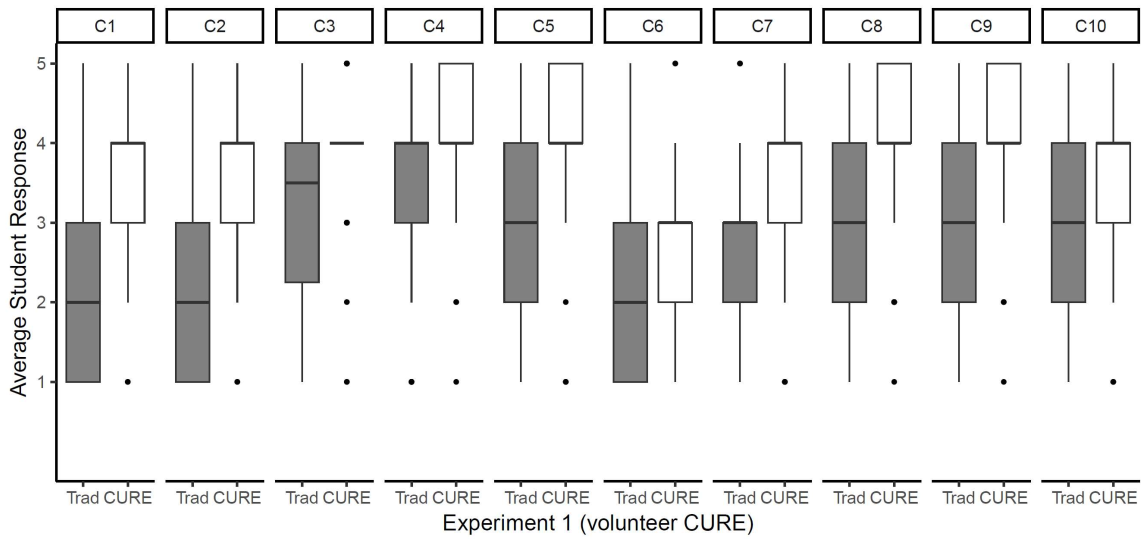


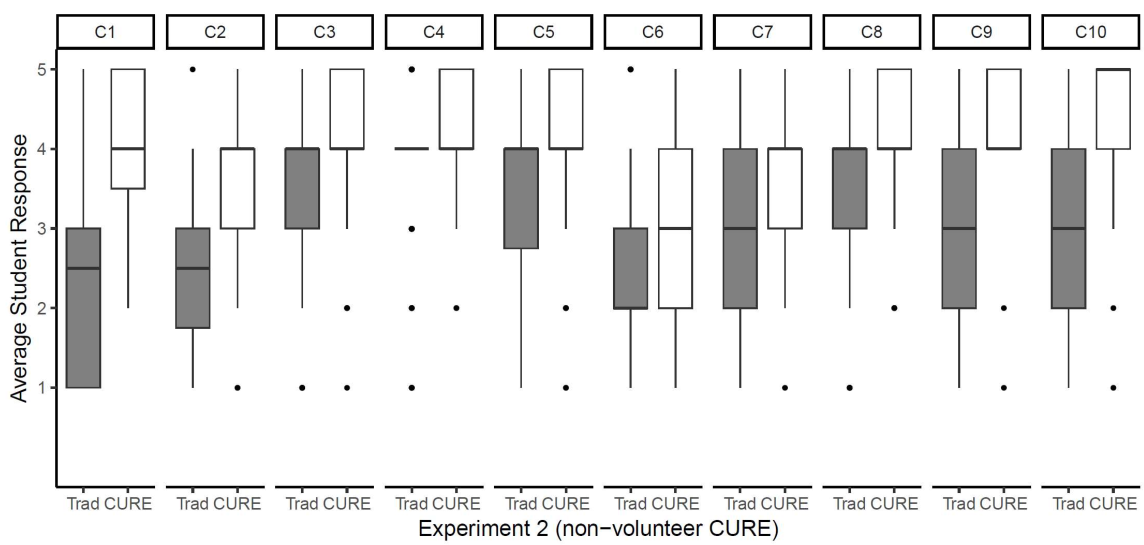
Results from individual items on the cognitive ownership sub-scale of the POS

Horizontal bars indicate median and boxes represent the interquartile range for each item. The items included on this subscale are:

(C1) My research will help to solve a problem in the world.
(C2) My findings were important to the scientific community.
(C3) I faced challenges that I managed to overcome in completing my research project.
(C4) I was responsible for the outcomes of my research.
(C5) The findings of my research project gave me a sense of personal achievement.
(C6) I had a personal reason for choosing the research project I worked on.

(C7) The research question I worked on was important to me.
(C8) In conducting my research project, I actively sought advice and assistance.
(C9) My research project was interesting.
(C10) My research project was exciting.





Regression results, lecture exams using ln(odds of exam score) as outcome variable

Call: Ln(odds of fraction of exam points) ~ Lab Type * Subgroup + Sex + ln(5-Concurrent GPA)

	Experiment #1 (Volunteer)			**Experiment #2 (Non-volunteer)**		
	Ln(exam odds)			**Ln(exam odds)**		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*

Intercept	1.82	1.67 – 1.97	**<0.001**	1.76	1.57 – 1.94	**<0.001**
CURE (ref = Trad)	0.23	-0.02 – 0.48	0.066	0.04	-0.16 – 0.24	0.708
Pell only (ref = neither Pell nor BIPOC)	0.04	-0.15 – 0.24	0.654	-0.08	-0.30 – 0.14	0.470
BIPOC only	0.13	-0.13 – 0.39	0.333	-0.01	-0.26 – 0.25	0.961
Both Pell elig and BIPIC	-0.14	-0.36 – 0.08	0.200	-0.13	-0.38 – 0.12	0.298
Female sex (ref = male)	-0.13	-0.27 – 0.02	0.085	-0.20	-0.35 – -0.05	**0.010**
ln(5-concurrent GPA)	-1.20	-1.37 – -1.03	**<0.001**	-1.15	-1.32 – -0.97	**<0.001**
CURE * Pell only	-0.15	-0.58 – 0.27	0.478	-0.12	-0.48 – 0.24	0.524
CURE * BIPOC only	-0.43	-0.99 – 0.13	0.130	-0.25	-0.65 – 0.16	0.235
CURE * Both Pell and BIPOC	-0.42	-0.85 – 0.00	**0.051**	-0.53	-0.93 – -0.14	**0.009**
Observations	432			406		
R^2 / R^2 adjusted	0.377 / 0.364			0.384 / 0.370		