Supplemental Material

CBE—Life Sciences Education

Meaders *et al.*

Supplemental Materials

Table of contents: **Pgs.**

Supplemental Appendix S1. Pilot Survey Details.

Below is the original question used in the Fall 2017 pilot survey, details about the student sample population, and details about how we incorporated feedback to revise this question.

Pilot Survey Questions

Question 1

On average, for what percent of class time do you expect the following to occur?
(slide bar with values from 0-100)

a. Students listen to lecture
b. Students work alone to answer clicker questions (questions that require students to enter their answers through a digital device such as a clicker, phone, or computer), worksheets, or other problems.
c. Students work in groups to answer clicker questions, worksheets, or other problems

This question was intended as a way to explore how students expected in-class time to be used.

Pilot Survey Student Sample

Student responses came from 2540 students taught by nine instructors in three different STEM subjects: biology, chemistry, and physics at two research-intensive institutions.

Pilot Survey Feedback

Feedback from undergraduate and graduate students, faculty, and staff helped us revise this question. We added a note to emphasize to students to consider the portion of their course that met during the specified lecture time, and we changed the language of the question to focus on instructor-led activities instead of student activities as a way to emphasize that we were asking about how class time is *meant* to be spent, instead of the commitment level of particular students to the activities. We also added an "other" option, and descriptions to each of the activities we asked students to identify. Finally, we decided to force student responses to each of the options to add to 100% of class time, because in the pilot survey we saw wide variation within student responses across all three of the options (for example, a student could select 100% for a, b, and c in question 1 of the pilot). The revised question is shown in Supplemental Appendix S2, Question 1.

Supplemental Appendix S2: Survey questions with instructions and notes that we provided for students in the fall 2018 survey.

Differences in wording between Question 1 on the First-Week and Mid-Semester Survey are highlighted in bold. For Question 1, as students filled in percentages, the total number at the bottom changed to match the total sum of each of the four options. Students were not able to move on from the question until their responses totaled 100%.

First-Week Predictions Survey Questions

Answer all of the following questions while thinking specifically about the portion of your [course #] course that takes place [days and times]. Please do NOT include any laboratory or recitation components of the course when answering these questions.

Question 1

Consider the portion of your current [course #] class that meets on [days and times]. On a typical day, for what percentage of class time **do you expect** the following to occur?

Make sure your answers total 100%

a. The instructor lectures to the students. *For example, the instructor presents material to the students while students are asked to listen and take notes*: ___
b. The instructor asks students to work alone. *For example, students are asked to answer clicker questions (questions that require students to share their answers through a digital device such as a clicker, phone, or computer, or through non digital means such as colored cards), complete worksheets, or solve other problems. Please do not include taking notes*. : _____
c. The instructor asks students to work in groups. *For example, students are asked to work in groups to answer clicker questions, complete worksheets, or solve other problems*. : _____
d. The instructor asks students to do other things. *For example, students are asked to watch a video or demonstration or to give presentations*. : _____

Total: _____

Question 2

What experiences or information did you use to make predictions about how class time will be spent (for example, experiences or information you received before or during the semester)?

Mid-Semester Perceptions Survey Question

Answer all of the following questions while thinking specifically about the portion of your [course #] course that takes place [days and times]. Please do NOT include any laboratory or recitation components of the course when answering these questions.

Question 1

Consider the portion of your current [course #] class that meets on [days and times]. On a typical day, for what percentage of class time **does** the following to occur?

Make sure your answers total 100%

a. The instructor lectures to the students. *For example, the instructor presents material to the students while students are asked to listen and take notes*: ___
b. The instructor asks students to work alone. *For example, students are asked to answer clicker questions (questions that require students to share their answers through a digital device such as a clicker, phone, or computer, or through non digital means such as colored cards), complete worksheets, or solve other problems. Please do not include taking notes*. : _____
c. The instructor asks students to work in groups. *For example, students are asked to work in groups to answer clicker questions, complete worksheets, or solve other problems*. : _____
d. The instructor asks students to do other things. *For example, students are asked to watch a video or demonstration or to give presentations*. : _____

Total: _____

Supplemental Appendix S3: First-week and mid-semester survey demographic questions

Note: You may choose to leave any or all of the following questions blank. Your answers will be used to better understand characteristics of students taking this survey.
Question 1: Is this your first-semester taking courses on a college campus? ● Yes ● No ● Prefer not to answer
Question 2: Are you a transfer student? ● Yes ● No ● Prefer not to answer
Question 3: Gender ● Male ● Female ● Not listed above _____ ● Prefer not to answer
Question 4: Race/Ethnicity (select all that apply)[1] ● American Indian or Alaska Native ● Asian ● Black or African American ● Hispanic or Latino ● Native Hawaiian ● White ● Not listed above _____ ● Prefer not to answer
Question 5: Did you speak English at home when you were growing up? ● Yes ● No ● Prefer not to answer
Question 6: Are you an international student? ● Yes ● No ● Prefer not to answer
Question 7: Highest level of education completed by at least one of your parents ● Did not complete high school ● High school/GED ● Some college (but did not complete college) ● Associate's degree (2-year degree) ● Bachelor's degree ● Master's degree ● Advanced graduate degree (e.g., DVM, MD, PhD) ● Unknown ● Prefer not to answer

[1] We asked students to select all that apply from the race/ethnicity categories, as defined by U.S. Census Bureau (2018). We categorized students who selected any of the following: Black or African American, American Indian or Alaska Native, and Hispanic or Latino as URM (National Center for Science and Engineering Statistics, 2017).

Supplemental Table S1: Comparison of survey population to course enrollment

The percent of students who responded to the first-week survey. Positive percent differences indicate that our survey had an overrepresentation of a demographic group's responses compared to the demographics of those enrolled in the course. Negative percent differences indicate that our survey had an underrepresentation of a demographic group's responses. Absolute differences greater than 15% are denoted in bold text.

Course	Male	Female	URM
1	3.6	-3.6	5.2
2	9.3	-9.3	8.0
3	1.9	-1.9	5.7
4	4.2	-4.2	9.7
5	-0.9	3.0	6.1
6	**-15.1**	14.7	7.8
7	4.9	-5.5	7.8
8	3.7	-3.7	-0.4
9	3.9	-3.9	2.3
10	2.10	-2.1	4.0
11	11.10	-11.1	3.0
12	3.9	-3.9	1.2
13	1.8	-1.8	3.9
14	7.1	-7.1	2.2
15	9.1	-9.0	2.2
16	9.3	-9.3	1.4
17	12.6	-12.6	-8.4
18	-3.5	3.5	-1.9
19	**17.0**	**-17.0**	-3.2
20	-5.4	5.4	1.1

Supplemental Table S2: One-way ANOVA results from students enrolled in multiple courses.

First Week Survey

Predictor	Sum of Squares	df	F value	*p* value
Student ID	35163	78	2.0728	6.58e-04 ***

*** $p < 0.001$.

Supplemental Appendix S4: Model selection using the complete dataset.

Model comparison for selection of random effects and fit statistics. The bolded model indicates the model with the lowest AIC and BIC values.

Model	AIC	BIC	Degrees of freedom
Model 1	14369.97	14429.39	11
Model 2	**14153.71**	**14223.93**	**13**
Model 3	14297.66	14367.87	13
Model 4	14152.83	14228.45	14
Model 5	14337.50	14402.31	12
Model 6	14155.96	14226.18	13
Model 7	14160.40	14230.62	13

Model 1: lm(Lecture ~ FirstGen + URM + FSCC + International + English + transfer + Gender + CourseSize, data=FirstWeek, REML=T)

Model 2: lmer(Lecture ~ FirstGen + URM + FSCC + International + CourseSize + English + transfer + Gender + (1|Instructor) + (1|ID), data= FirstWeek, REML=T)

Model 3: lmer(Lecture ~ FirstGen + URM + FSCC + International + CourseSize + English + transfer + Gender + (1|University/ID), data= FirstWeek, REML=T)

Model 4: lmer(Lecture ~ FirstGen + URM + FSCC + International + CourseSize + English + transfer + Gender + (1|University/Instructor) + (1|ID), data= FirstWeek, REML=T)

Model 5: lmer(Lecture ~ FirstGen + URM + FSCC + International + CourseSize + English + transfer + Gender + (1|ID), data= FirstWeek, REML=T)

Model 6: lmer(Lecture ~ FirstGen + URM + FSCC + International + CourseSize + English + transfer + Gender + (1|Course.Code) + (1|ID), data= FirstWeek, REML=T)

Model 7: lmer(Lecture ~ FirstGen + URM + FSCC + International + CourseSize + English + transfer + Gender + (1|University/Instructor), data= FirstWeek, REML=T)

Analysis of variance between Model 2 and Model 4, to confirm that these two models were equivalent.

Model	Degrees of freedom	AIC	BIC	Log Likelihood	Chi squared	*p* value
Model 2	13	14184	14254	-7078.9		
Model 4	14	14184	14260	-7078.0	1.83	0.18

Variance Inflation Factor test for multicollinearity between factors included in best fitting model. A value of 1 indicates no collinearity between a factor and the other predictors included in the model.

Predictor	**Variance Inflation Factor**	**Df**
Course Size	1.00	2
First-generation status	1.01	1
First-semester status	1.01	1

Analysis of variance between best fitting model and null model.

Model	**Degrees of freedom**	**AIC**	**BIC**	**Log Likelihood**	**Chi-square**	***p* value**
Final Model	8	14176	14219	-7079.9	43.967	6.518e-09 ***
Null Model	4	14212	14233	-7101.9		

*** $p < 0.001$

Final Model: lmer(Lecture ~ CourseSize + FirstGen + FSCC + (1 | Instructor) + (1|ID), data= FirstWeek, REML=T)

Null Model: lmer(Lecture ~ (1 | Instructor) + (1|ID), data= FirstWeek, REML=T)

Supplemental Appendix S5: Model selection using data set with only one response per student (single-response-per-student).

Demographic characteristics of the student responses for the First-Week survey, after randomly removing one response for each student enrolled in multiple courses (n = 1548 students). Total numbers within each group, and percentage out of the total number of responses are included.

Student Variables **College Experience**	**First-Week Survey**
First-Semester	732 (47%)
Returning Student	816 (53%)
English Spoken at Home	
English Spoken at Home as a Child	1400 (90%)
English Not Spoken at Home as a Child	148 (10%)
First-Generation status	
First-Generation	415 (27%)
Continuing Generation	1133 (73%)
Gender	
Male	762 (49%)
Female	786 (51%)
International Student	
Domestic	1460 (94%)
International	88 (6%)
Transfer student	
Non-Transfer	1384 (89%)
Transfer	164 (11%)
URM status	
URM	262 (17%)
non-URM	1286 (83%)

Course characteristics of the student responses for the First-Week survey, after randomly removing one response for each student enrolled in multiple courses (n = 1548 students). Total numbers within each group, and percentage out of the total number of responses are included.

Course Variables **Course Size**	**First-Week Survey**
Small (< 50 students): 3 sections	43 (3%)
Medium (51 - 110 students): 6 sections	216 (14%)
Large (> 110 students): 13 sections	1289 (83%)
Subject	
Biology	522 (34%)
Chemistry	180 (12%)
Computer Science	156 (10%)
Earth Science	47 (3%)
Economics	105 (7%)
Engineering	16 (1%)
Forestry	37 (2%)
Math	65 (4%)
Physics	203 (13%)
Statistics	217 (14%)
University	
1	817 (53%)
2	547 (35%)
3	184 (12%)

Model comparison for selection of random effects and fit statistics. The bolded models indicate the model with the lowest AIC values.

Model	**AIC**	**BIC**	**Degrees of freedom**
Model 1	13591.01	13649.81	11
Model 2	**13390.97**	**13455.10**	**12**
Model 3	13526.48	13590.62	12
Model 4	**13390.08**	**13459.56**	**13**

Model 1: lm(Lecture ~ FirstGen + URM + FSCC + International + English + transfer + Gender, data=FirstWeek, REML=T)

Model 2: lmer(Lecture ~ FirstGen + URM + FSCC + International + CourseSize + English + transfer + Gender + (1|Instructor), data= FirstWeek, REML=T)

Model 3: lmer(Lecture ~ FirstGen + URM + FSCC + International + CourseSize + English + transfer + Gender + (1|University), data= FirstWeek, REML=T)

Model 4: lmer(Lecture ~ FirstGen + URM + FSCC + International + CourseSize + English + transfer + Gender + (1|University/Instructor), data= FirstWeek, REML=T)

Analysis of variance between Model 2 and Model 4, to confirm that these two models were equivalent.

Model	Degrees of freedom	AIC	BIC	Log Likelihood	Chi squared	*p* value
Model 2	12	13421	13486	-6698.7		
Model 4	13	13422	13491	-6697.8	1.829	0.1762

***Best fitting model after model selection*.** The t value reported is the (regression coefficient)/(standard error).

Predictors	Estimate	Std. Error	t value	*p* value	2.5% Confidence Interval (CI)	97% Confidence Interval (CI)
(Intercept)	64.49	3.10	20.86	6.84e-15 ***	58.47	70.45
Course Size						
Medium Course	-20.64	4.94	-4.18	0.00052 ***	-29.98	-11.16
Small Course	-20.28	6.28	-3.23	0.00274 **	-31.55	-8.23
Continuing Generation	3.22	1.08	2.97	0.00307 **	1.10	5.35
Returning Student	4.43	0.98	4.5	7.38e-06 ***	2.68	6.35
Random effects						
Instructor	ICC = 0.38					
Observations: 1548 students $R^2m = 0.145$/ $R^2c = 0.343$						

* $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$

Variance Inflation Factor test for multicollinearity between factors included in best fitting model. A value of 1 indicates no collinearity between a factor and the other predictors included in the model.

Predictor	Variance Inflation Factor	Df
Course Size	1.00	2
First-generation status	1.01	1
First-semester status	1.01	1

Supplemental Appendix S6: Calculating the percent of lecture in two-minute time intervals

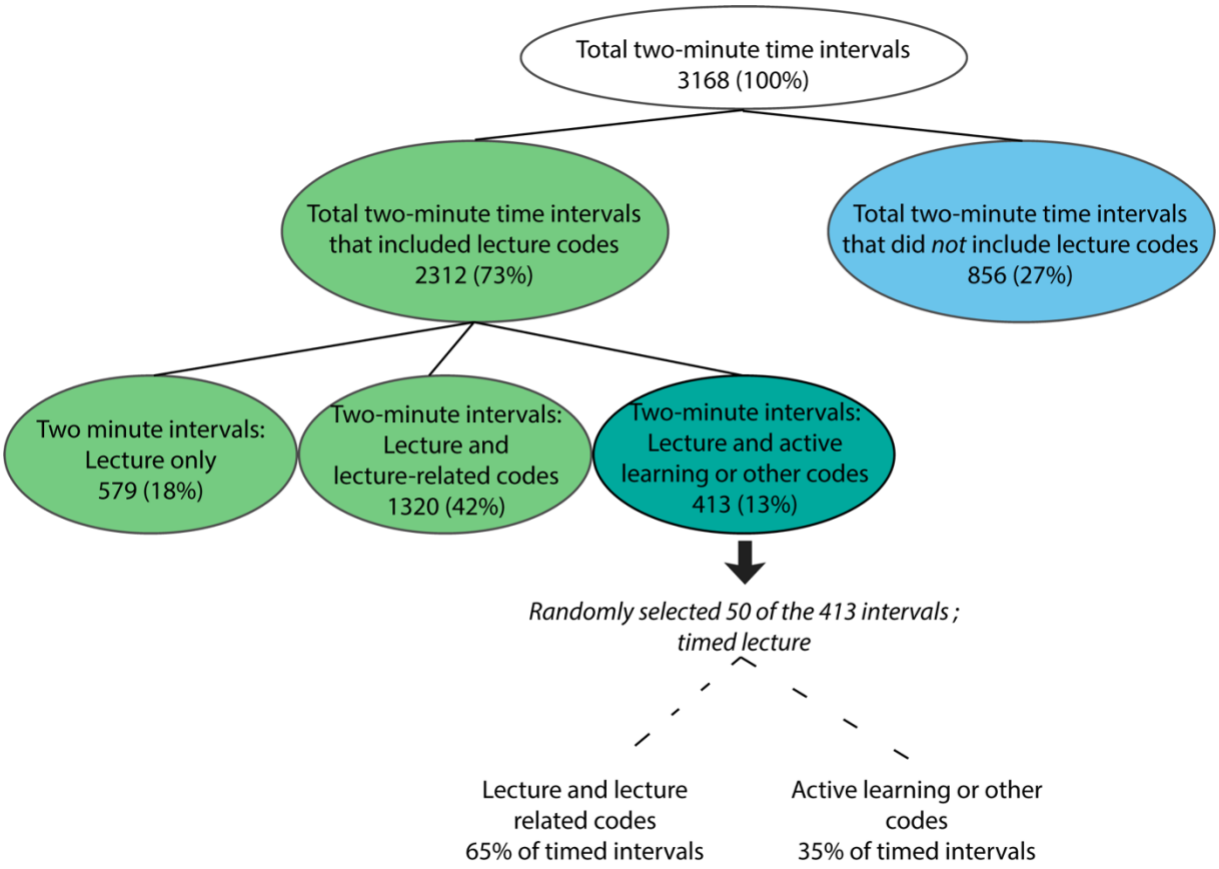
Analysis of average two-minute intervals with lecture across courses. As outlined in the methods, for each class period observed, we calculated the percent of two-minute time intervals that included the lecture code. We then calculated the average across all of the class periods observed, which is 74% and shown in Figure 2D.

Analysis of two-minute time intervals for all instructors and all class periods. The analysis above provides an overall estimate of how much lecture students experienced within a given course. We compared this estimate with how much lecture students predicted they would experience within a typical class period.

However, one caveat of the analysis is that the average percent of two-minute intervals that included lecture may overestimate the amount of lecture, because lecture and other active learning codes can be selected in the same two-minute interval.

In order to determine how many two-minute intervals may be contributing to an overestimation, we examined the frequency and co-occurrence of lecture and other codes, which is summarized below. For this analysis, the total number of two-minute intervals that included lecture across all courses was divided by the total number of two-minute intervals. This calculation results in the percentage of codes that included lecture overall (73%). This number differs by 1% from the average of percent of two-minute intervals within class periods depicted in Figure 2D because it is an overall mean rather than a grand mean of course means.

Co-occurrence of lecture and other COPUS codes. Green ovals indicate lecture only or lecture related codes; blue ovals indicate two-minute intervals without lecture related codes; the blue/green oval indicates two-minute intervals with both lecture and active learning or other codes. Results from timing lecture in a random sample that included both lecture and active learning or other codes are presented as plain text without shapes.



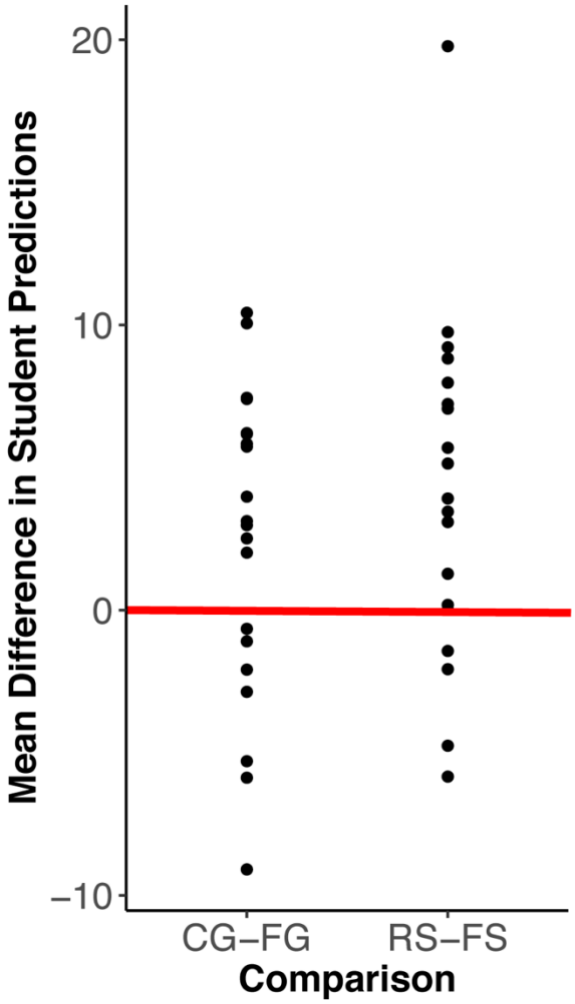
For the 73% of two-minute time intervals that included lecture, there are three possibilities. The two-minute time intervals include:
1. Lecture only: 18% of the total two-minute time intervals
2. Lecture and lecture-related codes (Real-time writing, Follow-Up, Posing non-clicker questions, Answering questions, Administration): 42% of the total two-minute time intervals
3. Lecture and active learning or other codes (Clicker questions, Moving & Guiding, 1 on 1 interactions, Other): 13% of the total two-minute time intervals

The first two possibilities described above do not overestimate lecture, as no active learning related or other codes were coded during the intervals. However, the third possibility could overestimate the amount of lecture. To estimate how much, we selected 50 of those intervals and used a stopwatch to time how much time was dedicated to lecture.

The timing analysis revealed that 65% of actual time from these two-minute intervals was dedicated to lecture or lecture related codes. Applying this to all of the two-minute intervals that included lecture and active learning or other codes equals 268.5 or 8.5% of the total codes. Instead of adding 18% + 42% + **13%** = 73% to tally the two-minute time intervals that include lecture, we can correct for the possibility of lecture, active learning, and other codes co-occurring and calculate 18% + 42% + **8.5%** = 68.5%.

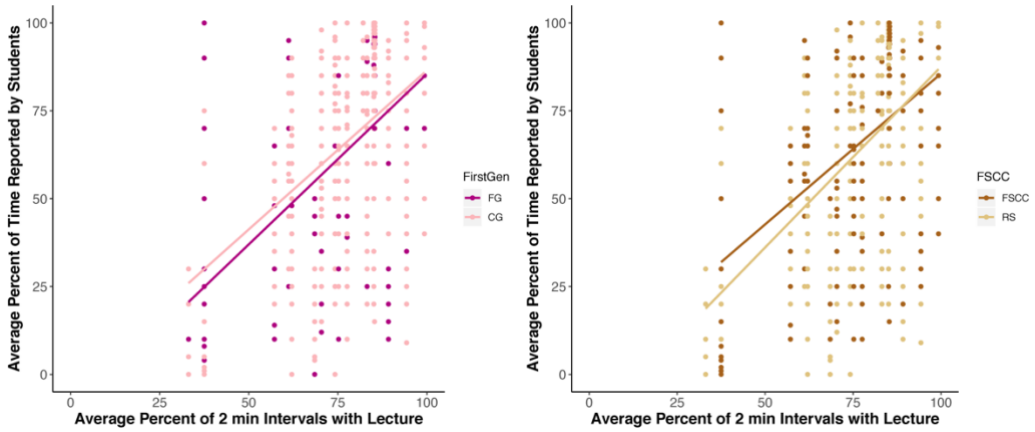
Supplemental Appendix S7: Differences in student predictions at the course-level.

Differences in student predictions based on demographic group



Plot displaying the difference in means between Continuing Generation (CG) – First Generation (FG) and between Returning Student (RS) – First-Semester (FS) student predictions of how much in-class time would be dedicated to lecture. Each dot represents the difference from one course. The red line at 0 indicates where there are no differences between the two demographic groups. Dots above the line (positive values) indicate that continuing generation or returning students predicted a greater amount of lecture than their peers, and dots below the line (negative values) indicate that first-generation or first-semester students predicted a greater amount of lecture than their peers. Across the majority of courses, continuing generation (13/20 courses) and returning students (15/19 courses) predicted more lecture than their peers. One course was excluded from comparing the difference in returning student and first-semester student predictions, as only returning students completed the survey.

Supplemental Appendix S8: Correlation between individual student reports and COPUS observation data, disaggregated by first-generation or first-semester student status.



Scatterplot of individual student reports of in-class time spent lecturing at the mid-semester point compared to the average observed percent of two-minute intervals that contained lecture for that course. Student responses were disaggregated by first-generation or continuing generation status, or by first-semester or returning student status. FG=First-generation, CG= continuing generation, FSCC=First-semester on a college campus, RS=returning student

Regression lines:
First generation: 0.97x - 11.54; $R^2 = 0.30$
Continuing generation: 0.90x - 4.00; $R^2 = 0.23$
First-semester on a college campus: 0.86x - 0.24; $R^2 = 0.24$
Returning Student 1.03x -15.32; $R^2 = 0.28$

Analysis of covariance summary statistics: First-generation status. The amount of lecture observed is a significant predictor of the amount of lecture reported by students, but first-generation status had borderline effects on student reported lecture.

Predictor	**F value**	**Df**	***p* value**
Amount of Lecture Observed via COPUS	428.55	1	< 2e-16 ***
First-Generation Status	3.96	1	0.05 N.S.
Type II ANOVA results were generated using the formula Student reported lecture ~ Observed lecture + First-generation status			

*** $p < 0.001$

Analysis of covariance summary statistics: First-semester status. The amount of lecture observed is a significant predictor of the amount of lecture reported by students, but first-semester status had no significant effects on student reported lecture.

Predictor	**F value**	**Df**	***p* value**
Amount of Lecture Observed via COPUS	442.05	1	< 2e-16 ***
First-Semester on a college campus	2.98	1	0.08 N.S.
Type II ANOVA results were generated using the formula Student reported lecture ~ Observed lecture + First-semester status			

*** $p < 0.001$