Appendix A4
Study Methods for Connecting Observation Data and Student Survey Data

A4.1 Introduction

Chapter 4 was intended to explore direct connections between the classroom observation data and the student survey outcomes, addressing the research question:

- How do student learning outcomes relate to the nature of the instruction they experience?

In other chapters, we have used the labels "IBL" and "non-IBL" to gloss the broadly different approaches used in particular classes, comparing student groups by these campus-designated labels. However, these labels hide a good bit of variation (see Section 2.3). In Chapter 4 we attempted to directly link the data from the classroom observations to student gains reported on the SALG-M post-survey. There we discussed the principles involved in linking the observation data to the student self-reported gains from the surveys. Here, we record the specific procedures used in sufficient detail that they could be reproduced in future studies.

A4.2 Study Sample

This analysis relies on two types of data—student post-course surveys and classroom observations—that must be collected from the same course sections in the same term. We collected observation data for 43 course sections at three campuses. The details of the observation study sample are available in Chapter 2 (Section 2.1) and Appendix A2. Unfortunately, we did not obtain post-surveys from seven of the sections included in this sample. The details of the survey study sample can be found in Chapter 3 (Section 3.1) and Appendix A3. Likewise, we could not gather observation data from all sections that returned student post-surveys. The final data set for the combined observation/survey analysis includes 30 IBL and six non-IBL sections, with averages representing 670 students. The Hierarchical Linear Model for full survey data set (see Section A4.5.2) includes 1239 student survey responses from 80 course sections.

A4.3 Data Collection

The data collection methods for the observation study are discussed in detail in Chapter 2 (Section 2.1) and Appendix A2. The data collection methods for the survey study are detailed in Chapter 3 (Section 3.1) and Appendix A3. The observation data and the survey gains data pertain to two different units of analysis: the classroom observation data describe course sections, and the survey data describe individual students. To address this mismatch, we computed student gain averages for each course section. In the next section, we detail how the specific observation data composite variables were constructed and labeled.

A4.4 Construction of Variables

We constructed several composite variables for the purpose of identifying student-centered approaches in classroom activities and practices.

A4.4.1 Classroom Time Variables

These are based on observers' minute-by-minute records of what activities were conducted during class and who led each activity, averaged over several hours of observation of multiple class sessions.

- Total percentage of student-centered time – indicates total percentage of class time spent on student-centered activities. This variable combines the percentage of class time spent on student presentation, percentage of time spent on discussion, percentage of class time spent on group work, and percentage of class time spent on computer-assisted learning.
- Total percentage of student-led time – indicates total percentage of class time spent with an individual student or a group of students in the leadership role. This variable combines the percentage of time with an individual student in the leadership role, percentage of class time with a group of students leading the class, and the percentage of time with the entire class in the leadership position.

A4.4.2 Observer Survey Ratings Variables

Observer ratings variables are based on the observer survey ratings of the frequency of specific student and instructor behaviors discussed in detail in Chapter 2 (Section 2.2.7) and Appendix A2. We constructed four composite variables from the observer survey items based on the four clusters of items (factors) that emerged from the Exploratory Factor Analysis. The resulting composite variables use mean observer ratings for these items as reported for several separately observed class sections; they are discussed in detail below.

- Student-instructor interaction: composite index that averages the ratings for seven items from the observer survey, indicating the frequency with which:
    - students offer ideas during class
    - students receive personal feedback on their work
    - instructors listen to students' ideas
    - instructors give concrete feedback on students' work
    - students ask questions
    - instructor offer help to students
    - inverse rating for instructors express their own ideas or solutions to problems.

For the first six items included in this composite variable, a higher rating indicates a higher frequency of various student-instructor interactions. However, the seventh item—instructors express their own ideas or solutions to problems—is just the opposite in that a higher rating for this item indicates lower interaction with students. Thus, in order to combine it with the other items in the student-instructor interaction composite variable, it was necessary to reverse the direction of this scale. Therefore, we subtracted the "instructors express their own ideas or

solutions to problems" ratings from 5, the highest rating possible, arriving thus at the rating for "instructors do not express their own ideas or solutions to problems."

- Student-student interactions: composite index that averages the mean frequency ratings for three items all related to students' interaction with each other:
    - students review or challenge others' work
    - students work together with others
    - students get help from others
- Students' role in setting course pace and direction: composite index that averages the mean frequency ratings for two items about course pace and direction:
    - student set pace or direction of the class time
    - inverse rating for instructors set pace or direction of the class time

Similarly to the student-instructor interaction, one item in this index has the opposite direction to the orientation of the index. While the first item in this scale points to more student contribution in setting the direction of the course, the second item indicates more instructor and thus less student contribution in these matters. Thus, we reversed the direction of the second item, by subtracting its rating from the highest rating of 5. This provided us with rating for "instructors do not set pace or direction of the class time." The average of both items points to students' active role in setting the pace and direction of the class.

- Instructor behaviors: composite index that averages ratings for two instructor behaviors:
    - instructors establish a positive atmosphere
    - instructors summarize or place class work in a broader context

A4.5 Data Analysis

A4.5.1 Correlation Analysis

To check for a quantifiable relationship between the classroom observation variables and the section means for student gains, we used the non-parametric Spearman correlation test. As most of our data in this study was not normally distributed, according to the Shapiro-Wilk test of normality, using a non-parametric correlation test is the most appropriate choice, since it is specifically suited for non-normally distributed data. We used SPSS (version 18) to compute the section means for student gains responses and then performed statistical analyses on the resulting course-level data.

We used scatter plot function in SPSS to obtain visual representations of the relationships between various classroom observation variables and section means for different student gains. We used the Add Fit Line option in the chart editor to experiment with linear, quadratic, and cubic approximations for the data and add them to the scatter plots. We also checked the statistical significance of linear, quadratic, and cubic approximations of the data by using the

Curve Estimation function in SPSS. Besides performing a statistical test, this function also provided the coefficient of determination (R^2) for each curve estimation, indicating the percentage of variability in the data that is explained by each curve.

A4.5.2 Hierarchical Linear Model for Full Survey Data Set.

We used a Hierarchical Linear Model to assess the relative impact of participation in IBL courses on self-reported student gains. Using this type of model assesses the relative influence of student characteristics such as gender and class year, and of course characteristics (including participation in the IBL program) on self-reported gains in the course.

The Hierarchical Linear Model follows the standard intercept model with student and course level variables. The model appropriate for analysis is described below. The general form of the equation is:

$$Y = \beta_{01} + \beta_1 X1 + \beta_2 X2 + \ldots + \beta_6 X5 + \varepsilon_i$$

$$\beta_{01} = \gamma_{00} + \gamma_{01} W_1 + \ldots + \gamma_{0k} W_k + \mu_{0i}$$

$$\beta_{11} = \gamma_{11} W_{1i}$$

Where β_{01} is the level 1 intercept, β_i are level 1 regression weights, X_i are level 1 variable scores, γ_{00} is the level 2 intercept, γ_l are level 2 regression weights, and W_k are level 2 variables scores. ε_i and μ_{0i} are error terms. Y is the outcome variable that is being predicted; in this case learning gains. Level 1 refers to student-level variables, and Level 2 to course-level variables.

We formed factor variables from items loading greater than 0.4 on Varimax rotated factors using a Principal Components extraction. Factors scores were then assessed for independence from each other: only relatively independent factors with less than 0.3 correlation between factors were used in the analysis. The composite variables with low internal reliability were also excluded from the model.

The outcome variable for the model is a composite variable of weighted factor scores for five learning gains items from the SALG-M post-survey, all related to cognitive gains in understanding mathematical concepts, thinking, and relationships: the main concepts explored in class, the relationships among the main concepts, students' own ways of mathematical thinking, how mathematicians think and work, and how ideas from this class relate to ideas outside mathematics. The composite variable showed high internal reliability at $\alpha = 0.88$. The outcome variable is defined such that it has a mean of 0 and Standard Deviation of 1.

Independent variables included student demographics variables of gender and college class level. Course level variables included participation in the IBL program and participation in pre-service teaching courses. The coding of the variables is detailed in Table A4.1.

Table A4.1: Student-Level and Course-Level Variables

Student Level	Coded
Class standing: First Year Sophomore Junior Senior Graduate	 No=0 Yes=1
Gender	Male = 1 Female = 2
Ethnicity	Hispanic = 1 Not Hispanic = 2

Course Level	Coded
IBL course	IBL = 1 Non-IBL = 0
Pre-service teacher course	Pre-service course = 1 Not a pre-service course = 0

The means and standard deviations for the outcome factor variable at different levels of the independent variables are presented in Table A4.2. Table A4.2 also includes sample sizes for different levels of the independent variables.

Table A4.2: Student Gains Factor Variable by Levels of Independent Variables

Independent Variables		Outcome Student Gains Factor Variable		
		Mean	Standard deviation	N
IBL	Non-IBL	-0.11	0.90	366
	IBL	0.05	1.03	873
	(Difference)	0.16		
Pre-service teacher course	No	0.10	0.91	953
	Yes	-0.35	1.18	286
	(Difference)	0.45		
Gender	Female	-0.13	1.05	647
	Male	0.14	0.92	575
	(Difference)	0.27		
Class year	First-year	0.11	0.89	438
	Sophomore	0.17	0.94	172
	Junior	-0.18	1.08	277
	Senior	-0.09	1.08	299
	Graduate	0.45	0.63	15

A4.6 *Hierarchical Linear Model for Combined Observation-Survey Data Set*

We constructed a second Hierarchical Linear Model with the intention to include the classroom observation variables. This model examines the effect of three course-level variables on student outcomes: total percentage of student-centered time, plus the previously tested indicators of IBL

and pre-service courses. This model does not use any student-level variables (such as gender and class year) as predictors of outcomes.

For the second model, the means and standard deviations for the outcome factor variable at different levels of categorical independent variables are presented in Table A4.3. The table also includes the sample sizes for various levels of the independent variables.

Table A4.3: Student Gains Factor Variable by Levels of Categorical Independent Variables

Student Level Independent Variables		Student Gains Outcome Factor Variable		
		Mean	Standard Deviation	N
IBL	Non-IBL	-0.25	0.96	177
	IBL	-0.08	1.02	493
	(Difference)	0.17		
Pre-service teacher course	Not Pre-service	-0.03	0.91	428
	Pre-service	-0.31	1.13	242
	(Difference)	0.29		

On the other hand, Table A4.4 includes descriptive statistics for the continuous independent variable in this model: total percentage of student-centered time.

Table A4.4: Descriptive Statistics for Percentage of Student-Centered Time

Course-Level Independent Variable	Mean	Standard Deviation	N
Percentage of class time spent on student-centered activities	51.42	28.43	39

A4.7 Conclusions and Limitations

This type of analysis requires a very large volume of data, collecting which is a very labor-intensive task. Even with over 1200 student survey responses and 300 hours of observation, the size of the data set is on the very edge of what is needed to extract good correlations or to construct a complete, two-level hierarchical model. The cost and effort required to document these linkages are very high and should be seriously considered when undertaking such type of analysis.