Geri Kerstiens
May 15, 2017
Graduate Student Assessment Fellows Program
Project Summary

My project involves the design and assessment of a new curriculum for the introductory chemistry laboratory courses at UC Berkeley. The team working on this project includes Prof. Anne Baranger, the director of undergraduate chemistry and PI for this project; Dr. Michelle Douskey, a senior lecturer for general chemistry courses in the chemistry department; Dr. MaryAnn Robak, a lecturer for organic chemistry courses in the chemistry department; Laura Armstrong, a graduate student in the SESAME program; myself, also a graduate student in SESAME; and various undergraduate volunteers over the course of the project. The project began its current iteration in 2012 with a grant from DOW to renovate the physical laboratory space for the introductory lab courses and to update the curriculum to include a focus on green chemistry topics and green lab practices. This project is very large in scope and involves many goals. For this program I chose to focus on the assessment of a particular outcome of the curriculum, namely the amount of student learning on the topic of scientific models.

The curriculum is composed largely of the lab manual, which includes background information and the experimental procedures the students will go through each week. It is implemented by the course instructor, who varies each semester, and the GSIs, who are in the labs giving instructions and setting the tone during the experiments. The experiments that make up the lab curriculum are first developed and tested by the lecturers, graduate students, and undergraduate volunteers on the project. This is done by reviewing current chemistry education and research literature, determining what topics should be covered in the relevant courses, and adapting the experimental procedures we find to the constraints of a three-hour introductory laboratory class period. The experiments are then implemented in the appropriate courses, and feedback on this implementation is collected (via the students, GSIs, and instructors) and used to refine the experiments before their next implementation. Assessment of the project is done by the PI, lecturers, and graduate students on the project, largely via weekly meetings where the progress of the project is discussed. The stakeholders for this project include DOW and the UC Berkeley chemistry department. The department is informed of the project's progress and outcomes through talks and meetings involving the project PI. Assessment data on the project is often a major focus of these events.

There were several motivations for assessing this project. First, as chemistry educators, we are interested in determining the effectiveness of our new curriculum: are we achieving what we hope to achieve with regards to student learning and attitudes toward chemistry? Second, as chemistry education researchers, we hope to publish materials on our curriculum for both the chemistry and science education communities. With regards to the assessment I conducted this semester, we wanted to know particularly how well our curriculum addresses students' knowledge of the nature of science (i.e. what is science, what does it mean to do science, how do we acquire scientific knowledge, etc.). Do the students show any gains in their understanding of these topics? Can they apply them in their laboratory activities?

Assessing student outcomes can involve a number of data sources. We have an extensive survey, given both before and after the course, to learn about students' confidence in the various course topics, their attitudes toward chemistry, their feelings about the course, and their use of available resources. We have conducted interviews to probe these survey topics in more depth.

Observations and video from inside laboratories provide another element to understanding the implementation of the curriculum and how changes in students' understanding may be initiated. Finally, students' work, such as their pre-lab exercises and lab reports, have been collected. These data sources can be used to assess a number of questions about student outcomes. The assessment for this project involved probing a particular outcome related to students' understanding of the nature of science: did the students make any gains in their understanding and use of scientific models, and what influenced those gains? Answering these questions involved looking at both observation data and students' report sheets, which they turn in at the conclusion of each week's experiment.

Observations were conducted in two lab sections of Chem 1AL (the introductory chemistry laboratory course for non-majors at UC Berkeley) during each lab period (with one three-hour lab period occurring per week) of the Fall 2015 semester by myself and five undergraduates. Students' report sheets were collected for each lab period, as well. The observation data is used to examine the implementation of the curriculum: how did GSIs approach the material, and how did students receive it as they conducted the experiments? The report sheets give us information on what the students have learned. Each report sheet has a number of questions probing students' understanding of new chemistry concepts, laboratory techniques, and the components of experimental design and execution. There are usually 1 to 3 questions each week that are meant to specifically target their nature of science knowledge. This assessment was confined to one particular experiment involving the composition of various polymers. One goal of this experiment is for students to learn more about the process of experimental design. In one question we ask them to create a chemical model that they will test with a given experiment. To assess this learning goal, I reviewed students' answers to this model question and assigned a score of 0 to 3 to indicate the level of sophistication of their models (see the appendix for a depiction of the question and coding scheme).

In reviewing the data, an interesting pattern emerged. One of the two sections performed significantly better on this question (see the appendix for the breakdown of students' scores by section). Because the materials for each section were identical, it must have been the implementation of the curriculum that led to this difference. In fact, the observation data details that the section with the higher scores had a GSI that spent much more time engaged in discussion with the class on the topic of scientific modeling. The other GSI gave only a short pre-lab lecture that spent little time on topics of modeling or experimental design. This was a trend throughout the semester. Of course, an alternate hypothesis to explain the difference in scores by section would be that the populations of the two sections were significantly different. Perhaps the higher scoring section had more experience with experimental design labs in the past. Perhaps they had more chemistry experience altogether. Their prior knowledge of chemistry topics could have been more extensive, or the class was composed of "better chemists," generally speaking. Demographic data from our survey is included in the appendix, and it shows that we can discount almost all of these explanations. (The higher scoring section did score slightly better with regards to their previous experience with experimental design.)

This data leads to an important conclusion: the curriculum is effective at improving students' understanding of scientific modeling, but this improvement is contingent on the implementation of the curriculum. The teaching choices of the GSIs seem to have a major impact in terms of the effectiveness of the curriculum. It is important to note, however, that this finding is preliminary. This data looked at a small subset of the Chem 1AL population during one experiment of one semester. A larger sample size and reproducibility are needed to make more substantial claims about our curriculum. With that in mind, more data was collected during the Spring 2017 semester. We took video of the GSIs in almost every section (approximately 10 to 15 sections have full data sets out of 19 sections) so we would have a record of what they say and do during their lab sections. We also have the students' report sheets so that we can perform the same coding of the model question. However, with the video we will be able to link the students' performance on the question with a more quantitative representation of the teaching methods of the GSIs. We also have other report sheet questions that examine students' understanding of modeling that must be analyzed to see if the pattern holds for other experiments throughout the semester. However, the early results are still useful in terms of allowing us to address potential shortcomings in our curriculum. We plan to assemble useful teaching tools from this data that could help the GSIs to more effectively implement the curriculum. They can see what methods are useful for approaching these somewhat daunting topics. In the end an implementation guide for the course instructor and GSIs would be the best outcome of this assessment.

Whenever a new curriculum needs to be assessed, it is most helpful to start with the motivations for the change. Why are you changing your curriculum? What do you hope to achieve with your efforts? Composing a clear list of your goals even before you begin designing your curriculum is actually the best practice, and it makes assessment of your curriculum much more straightforward. To assess your goals, you have to consider how each one is being addressed by the curriculum. If you think of all possible sources for the change you are hoping to see, you can more accurately determine if your curriculum is contributing to the outcome. Most importantly, you have to consider the implementation of the curriculum. The materials alone do not account for everything you see happening with the students. When assessing student outcomes, we have found it helpful to examine our materials to determine where students might be learning about these topics and where they can demonstrate their knowledge of them. Sometimes, it is within the course that this happens. Other times, we have to investigate these outcomes with surveys or interviews. Chemistry concepts are often found in report sheet questions. Attitude toward the discipline of chemistry, however, is easier to determine with a survey. Finally, curriculum design and assessment work best in tandem. Assessment is an important tool in determining where your curriculum is doing well and where it needs adjustment. Curriculum design should be iterative; you should expect it to take several cycles of implementation and refinement in order to reach your ultimate goals. Because of that time commitment, it is best to keep the scope of your goals reasonable. It is easy to gather a large amount of data; analyzing it is another story. Try to map your data collection plans onto your curricular goals to ensure you collect only useful data that you plan to examine.

Appendix 1: Model Question

Model 1:

Hypothesis: As the mass ratio of PVA to guar gum decreases, ______________________________
______________________________________.

Experiment:

Data:

Trends:

Conclusion:

Model 2: (not needed if conclusion confirms hypothesis)

Model 1 is the portion of this question that was analyzed to determine students' level of understanding of the construction of chemical models.

Appendix 2: Coding Scheme and Examples of Model Question Coding

Code for Model 1	Answer Content
0	Not a model
1	No connection to molecular properties
2	Vague, possibly untestable reference to molecular properties
3	Links macroscopic result with molecular level properties; explanation at the molecular level

Coding scheme for Model 1 question. The definition of "model" here is meant to pertain to a chemical model, something that attempts to explain a macroscopic phenomenon with molecular-level reasoning.

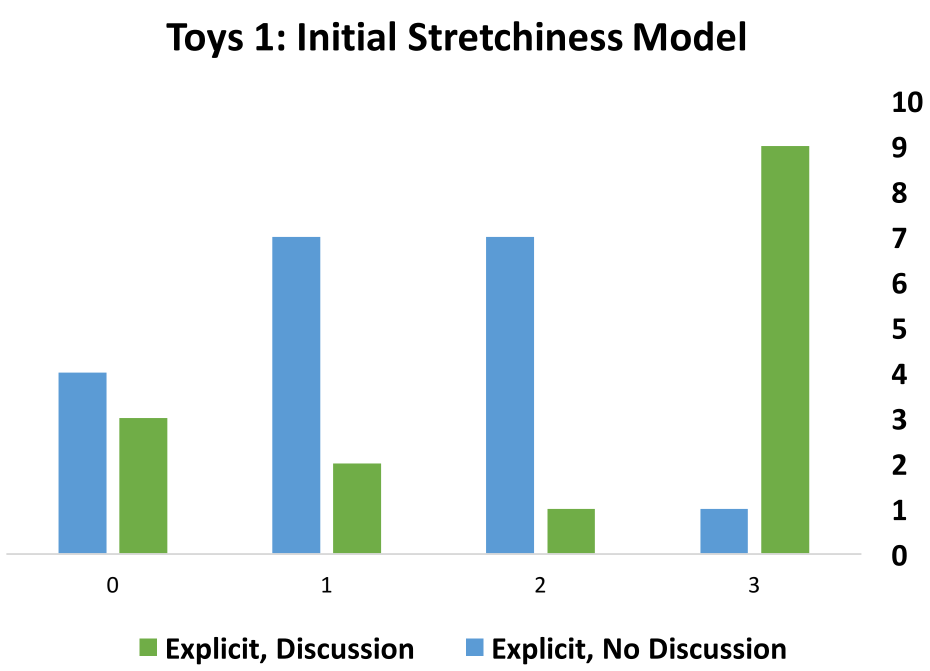
Code for Model 1	Example Answers
0	Stretchiness being tested as an average of gravitational fall over time for 3 trials.
1	Mixing different amounts (ratios) of PVA and guar gum will result in different stretchiness.
2	Simpler repeating units and increasing borax produces thicker substances.
3	Copolymerization with smaller, more lattice-like monomers yields more solid-like copolymers as a result.

These are example student answers that were coded according to the scheme presented in the above table.

Appendix 3: Data Analysis for Model 1 Question

	Explicit, Discussion	**Explicit, No Discussion**
Assigned Score	Number of Students	Number of Students
3	9	1
2	1	7
1	2	7
0	3	4
Average Score	**2.07**	**1.26**
p-value	**0.0485**	

These are the results of the Model 1 question coding from the two observed sections of Chem 1AL during Fall 2015. "Explicit, Discussion" is the designation for the section with the GSI who dedicated more time to discussing scientific modeling, while "Explicit, No Discussion" designates the other section. The p-value is less than 0.05, indicating that the average scores of the two sections are statistically significantly different.



These are the same results from the above table, shown graphically. The x-axis indicates the coding score, and the y-axis indicates number of students.

Section	Average Final Scores	P-Values
Explicit, Discussion	79.9	0.00706
Explicit, No Discussion	85.7	0.0867
Whole Class	88.4	--

Average scores on the Final Exam for the two observed sections and the entire Chem 1AL class. Note that the Explicit, Discussion section did not perform better than the Explicit, No Discussion section, indicating that a greater general talent for chemistry lab is most likely not the reason for that section's higher average score on the modeling question.

	Class	E, D	E, ND
Did not do labs	3.79	0	7.14
Only followed a given procedure	55.17	62.5	71.43
Sometimes designed a procedure or scientific experiment	41.03	37.5	21.43

This table shows the percent of students in the class and in the two observed sections who have not done labs in previous chemistry courses, only followed a given procedure in previous chemistry courses, and sometimes designed a procedure or experiment in previous chemistry courses. This represents the familiarity of the students in the course with experimental design. The Explicit, Discussion section did have a larger percentage of students with design experience than the Explicit, No Discussion section, but it did not have a larger percentage of students with design experience than the class as a whole.

	Class	E, D	E, ND
Mean	2.68	2.25	2.58

This chart shows the average number of semesters of chemistry students in the class and in the two sections have taken prior to this Chem 1AL course. Note that the Explicit, Disucssion section has, on average, taken fewer semesters of chemistry than the Explicit, No Discussion section, meaning prior chemistry experience is most likely not a major factor in the coding scores of the two sections.