The Impact of Nature of Science Instruction on the Chemistry Laboratory Experience

by

Geri Anne Kerstiens

A dissertation submitted in partial satisfaction of the requirements for the degree of Doctor of Philosophy in Science and Mathematics Education in the Graduate Division of the University of California, Berkeley

Committee in charge:

Professor Anne Baranger, Co-Chair
Professor Marcia Linn, Co-Chair
Professor Terry Johnson

Spring 2019
The Impact of Nature of Science Instruction on the Chemistry Laboratory Experience

© 2019

by

Geri Anne Kerstiens
Abstract

The Impact of Nature of Science Instruction on the Chemistry Laboratory Experience

by

Geri Anne Kerstiens

Doctor of Philosophy in Science and Mathematics Education

University of California, Berkeley

Professor Anne Baranger, Co-Chair

Professor Marcia Linn, Co-Chair

Recently, there have been many calls for an increase in instruction on the nature of science (NOS) in schools (i.e. NRC, 1996; NGSS Lead States, 2013). These calls recognize the importance of this topic at all levels of science education, but there is little guidance in terms of how to address it effectively in curricula. Similarly, there have been calls for reforms to chemistry laboratory courses (Hofstein & Lunetta, 2004). These reforms include an increase in the level of authenticity in laboratory activities along with a need to help students find a connection to these types of courses. This dissertation describes a project that attempts to address each of these issues with the creation of a general chemistry laboratory curriculum designed to include explicit instruction on topics from the nature of science.

The overall hypothesis for this dissertation is that including NOS topics (which fall into the categories of the nature of scientific knowledge, scientific practice, and the social and cultural aspects of science) in a laboratory curriculum could provide positive outcomes for students. For example, explicitly teaching students about the nature of scientific practice and how it impacts scientific knowledge, in the context of a practical laboratory course, could help them connect their own work in the course to larger scientific principles. This would enable them to apply this learning to future science courses or even outside of school in the real world, thus providing a path to meaningful learning and making the laboratory course a more positive experience for students. Additionally, explicit discussion of the nature of scientific practice in a laboratory course, where authenticity can be constrained by resources, time, and students’ lack of scientific experience, could alleviate some of these difficulties by allowing students to see how their activities, whatever they may be, relate to different aspects of authentic practice. Perhaps it could also help students undertaking high-level inquiry activities, such as experimental design, in an introductory course by providing a level of scaffolding for them to see how their actions in the lab are reflective of authentic scientific practice.
This dissertation describes and assesses a general chemistry laboratory curriculum that includes explicit instruction on the nature of science embedded in the context of the laboratory activities. The design process involved examining the already-established laboratory activities to determine what elements of NOS were implicitly present and then changing the curriculum and its implementation to explicitly address those topics. This was done iteratively, and data was collected from two different semesters in an attempt to determine the most effective presentation of NOS in the materials, along with the best instructional methods for impacting students’ understanding and application of these topics. While NOS topics of all kinds were included over the course of the semester, only the topics of scientific models and model building were specifically examined in the studies presented here. These topics were covered in the most depth during the course, and they represent the goals of the curricular redesign well: addressing topics of scientific practice as they relate to scientific knowledge in the context of practical scientific activities. Analyzing this subset of data still yields important determinations of the effectiveness of this curriculum.

It was determined from the three studies presented here that students did indeed make gains in their understanding of scientific models. They were also able to apply that understanding in writing their own models during laboratory activities. However, students were able to write the best models when they had an instructor who explicitly addressed these issues during lecture and directed the graduate student instructors (GSIs), who taught individual lab sections, to do the same. That semester also included a more streamlined set of materials concerning scientific models, more directly addressing the chemistry content of those labs in terms of developing models and designing experiments to test those models.

Finally, students were able to design more thoughtful and logical experiments to test their models when they encountered the curriculum with more direct prompts for considering the relationship between experiments and models along with an instructor who better articulated the role of experiments in model building. Thus, explicit instruction on scientific models, coupled with an adequate curriculum and engaged instructors, led to changes in students’ practices in the lab. Students were very receptive of NOS topics in the course, as well, as indicated by their responses on course evaluation questions on a class-wide survey. They found the labs enjoyable and pointed to NOS topics as some of the most valuable takeaways from the semester. Overall, then, it appears an established curriculum that has been updated to include explicit instruction on models and model building can lead to positive gains and positive experiences for students in a general chemistry laboratory course.
# Tables of Contents

Abstract .................................................................................................................................................. 1

Table of Contents ................................................................................................................................... i

Acknowledgements ................................................................................................................................. ii

Chapter 1: Introduction .......................................................................................................................... 1

Chapter 2: Theoretical Framework and Literature Review ................................................................. 4

Chapter 3: Curriculum Development and Assessment ........................................................................ 10

Chapter 4: The Positive Impacts of Explicit Instruction on Scientific Models in a Laboratory Curriculum ................................................................................................................................. 17

Chapter 5: How the Implementation of a Curriculum with Explicit Instruction on Scientific Models Can Impact Students’ Learning in the Lab ......................................................................................................................... 32

Chapter 6: The Impact of Explicit Instruction on Scientific Models on Students’ Scientific Practices ......................................................................................................................................................... 49

Chapter 7: Conclusions and Future Work ............................................................................................. 62

References ................................................................................................................................................. 64

Appendix I: Complete Fall 2017 Versions of the Polymers 1 and Polymers 2 Labs ......................... 67

Appendix II: HLM Assumptions ............................................................................................................. 97
Acknowledgements

I want to thank my family first because there is no way I could have completed this work without their support. To my husband Craig, for understanding and accepting my interest in this work and for providing incredible amounts of emotional support throughout this journey. And also for our wine collection. To my daughter Adeline, for her unconditional love, her understanding, and her help in taking care of her brother. To my son Evan, for his sweet cuddles and the sleep he sometimes lets me have. To my parents, Pat and Caroyla, for showing me that teamwork can result in a PhD, and for always supporting my love of learning and my need to do things the hard way. To my siblings, Trish, Sean, and Ben, for being the most loving and supporting brothers and sister anyone could ask for. I love you all deeply, and this achievement wouldn’t mean anything without you.

I want to thank my committee: my adviser, Anne Baranger, for her support and guidance during this long journey (and her delicious jam); Marcia Linn, for helping to make this dissertation better with each set of notes; and Terry Johnson, without whom there would be no curriculum to assess. Thank you for your attention to detail and your words of wisdom. And your handkerchief.

A special thank you to Michelle Douskey, for being there for me from the beginning as a mentor for teaching, a teammate in research, and a source of comfort in times of stress. To Angie Stacy, for guiding me to the SESAME program in the first place. To MaryAnn Robak and Laura Armstrong, for the many, many hours of work you put into this curriculum and for helping me make my piece of it a reality. To my research groupmates over the years, for listening to talks on my work semester after semester and always providing critical feedback and a place to vent. To Beatriz Brando, my dissertation buddy who felt my stress as her own and held my hand through this process. My work is better because of you all, and my life is richer for having known you.

Finally, I would like to acknowledge Dow and Cal-EPA for funding, without which there would be no project to analyze.
Chapter 1. Introduction

There is a growing consensus in the science education community that instruction on the nature of science (NOS) is a necessary part of science courses (NGSS Lead States, 2013; NRC, 1996). It could potentially provide benefits in the areas of student motivation, understanding of scientific content, and scientific literacy (Driver, Leach, Millar, & Scott, 1996). It is also important in educating the next generation of scientists, both in an effort to generate interest in such a career as well as in preparing students to be thoughtful, ethical practitioners of science. What role this type of instruction should take is dependent on the goals of the course. For example, there are a number of curricula that have been developed for non-scientists to help them become scientifically literate citizens (Schwartz, Lederman, & Crawford, 2004). There are courses designed for high school students in science courses as a means to help them learn about NOS concomitantly with the pertinent science content (i.e. Eastwood et al., 2012). However, few studies focus on the inclusion of NOS instruction in post-secondary science courses, and fewer still examine NOS education in a laboratory setting, though many suggest this is an area ripe for NOS learning to occur (Lunetta, Hofstein, & Clough, 2007).

What makes a chemistry laboratory course an interesting candidate for NOS instruction? The teaching laboratory is believed by many to offer unique educational opportunities to science students by allowing them to engage with discipline-specific topics in hands-on activities. Ultimately, this should give them a better idea of how science is practiced by actual researchers. However, there is little evidence that this is the case. The majority of chemistry laboratory courses offer students “cookbook” activities that require them simply to follow procedures without engaging them in the tenets of true scientific inquiry (Hofstein & Lunetta, 2004). Additionally, Hofstein and Lunetta note that students’ attitudes toward the teaching lab are largely negative (2004). They see it as having little value as part of their science education. There has been a call in recent years to improve this by increasing the level of inquiry (i.e. student-directed work) in lab activities (Singer, Hilton, & Schwingruber, 2005). However, this is logistically difficult for both large lab courses, due to the scarcity of equipment, and introductory lab courses, due to the lack of content and research knowledge of the students. There is also substantial evidence that inquiry activities alone are not enough to teach students about the broader tenets of science (Schwartz, Lederman, & Crawford, 2004).

NOS instruction, though not particularly prevalent in laboratory courses thus far, could be a way to “fill in the gaps” of a more traditional teaching laboratory model. By instructing students on the broader tenets of scientific practice, it is possible to help them see how their laboratory activities, despite being largely directed by an instructor and bearing little resemblance to authentic, self-directed research, are related to the bigger picture of scientific practice. It could help them understand that the laboratory exercises in which they engage have a broader application than simply leading them to a “right” answer or cementing their understanding of discipline-specific content that they were first exposed to in a lecture course. It is a chance for them to experience the various moving parts of authentic scientific research in a controlled environment, and it has the potential to improve students’ impressions of the teaching lab as a place where they can truly learn about science in a practical manner, rather than as a requirement that seems disconnected from their science education experience.
How, then, can a laboratory curriculum that effectively includes NOS instruction be created? The literature suggests that growth in NOS understanding can only come through explicit instruction on the topic, rather than simple participation in inquiry activities (Schwartz, Lederman, & Crawford, 2004). In fact, inquiry-based laboratory activities that include explicit instruction on NOS have had positive impacts on students’ NOS views (Yacoubian & BouJaoude, 2010). Contextualizing the NOS instruction, such as including it in a laboratory activity, has also had positive results (Eastwood et al., 2012). Finally, there must be a point of reflection for students to engage in concerning NOS (Schwartz, Lederman, & Crawford, 2004). A lab curriculum that successfully improves students’ understanding of NOS must follow these guidelines.

The chapters that follow detail the design and assessment of just such a laboratory curriculum for the general chemistry laboratory course at UC Berkeley. Its creation involved an examination of the existing curriculum to determine which elements of NOS were already implicitly present in the materials. These were then updated, along with suggestions for instruction methods, to explicitly discuss these ideas around the nature of scientific knowledge, inquiry, and societal impacts of scientific practice. Students were prompted to consider these NOS topics in the manual—throughout the background reading they did for each lab as well as during the execution of the actual experiments—as well as through their instructors during the experiment itself, in many cases. This was all in an attempt to help students learn about these topics in order to apply them in their lab practices, both within the course and afterward in future science classes or research.

The NOS topics covered by the curriculum range from the importance of argumentation at all stages of scientific practice to the role that a scientific community plays in establishing what is considered validated scientific knowledge. One topic that was covered in some detail was that of scientific models. The curriculum describes scientific models (in chemistry, in particular) as microscopic-level explanations of macroscopic phenomena. They are a particularly important component of scientific inquiry, and as such, they show up repeatedly in the course. There are four different labs that discuss the role of models in explaining the natural world and how scientists develop hypotheses and experiments to test and refine these models in order to add to or improve upon existing scientific knowledge. Essentially, scientific models and model building were presented as indispensable parts of scientific inquiry. As one of the main goals of this curriculum redesign was to improve the authenticity of labs and help students better understand how to participate in scientific practice, assessing their understanding and use of scientific models was an important goal for the research team. Therefore, the main points of focus of this dissertation are to determine if students made learning gains concerning scientific models, what instructional and materials choices were most effective for achieving these gains, and how this type of explicit instruction on scientific models could impact students’ practices in the lab.

Chapter 2 details the theoretical framework behind the NOS updates to the curriculum. It goes into some depth about the definition of NOS that was used to guide the development of the materials and how that definition was chosen. Chapter 3 describes the curriculum that was developed and the NOS topics what were covered in each activity. It also poses the research questions that are addressed in Chapters 4, 5, and 6. Chapter 4 seeks to determine what gains were made by students concerning scientific models and how much they valued this type of instruction. Chapter 5 is a quantitative analysis of students’ abilities to apply this knowledge and
write their own scientific models. It examines what types of instruction were most effective for helping students write better models. Finally, Chapter 6 examines students’ practices in the laboratory as they are asked to develop their own experiments to refine initial models concerning polymer solution properties in a high-inquiry activity. All of these studies yield important findings about how an established curriculum updated to include explicit NOS instruction can lead to gains in NOS learning and improvements in students’ experimental design choices.
Chapter 2. Theoretical Framework and Literature Review

The State of Science Education and the Teaching Laboratory

Traditional science education has often relied on teacher lectures and what Lemke calls the Triadic Dialogue: question-answer-evaluate (Lemke, 1990). This format often leads students to believe that science is a collection of de-contextualized facts that are meant to be memorized. They get no clear picture of what science actually is and how it works. They are not exposed to what it means to be a scientist. It is not surprising then that most in the research world are agreed that traditional methods do not work. One way the science education community wants to fix this is by increasing the amount of inquiry in science classrooms (Singer, Hilton, & Schwingruber, 2005).

What does inquiry mean? It is a term that appears often in the science education literature, and therefore, it can have many different uses. The NRC gives a useful overview (1996). Inquiry involves reviewing existing literature, determining a research question, constructing a hypothesis, developing and conducting an experiment(s), analyzing data, and making a determination of some kind (which could involve revising an initial hypothesis or developing a new experiment). When used in the context of science education, there can be various levels of scientific inquiry. Very rarely do students get to go through the entire process of inquiry as experienced by a professional scientist (this would be called authentic inquiry or high-level inquiry). More commonly, students go through pieces of the inquiry experience, and their level of inquiry is determined by how many of those pieces they are able to experience firsthand. Thus, a major goal of science education research is to find what level of inquiry is necessary and sufficient for producing some level of understanding (be it of scientific content or of the scientific process itself). The consensus of most of the literature is that inquiry is an important part of science education, and simultaneously, there is not enough of it (Singer, Hilton, & Schwingruber, 2005). Therefore, another goal of science education researchers is to find ways to increase the level of inquiry in students’ science education experiences.

One of these ways is through the teaching laboratory. The teaching lab has been a part of science education (in almost every domain) for over a century (Hofstein and Lunetta, 2004). It is a particularly memorable part of chemistry education for most former students. Most educators seem to agree that the teaching lab is a valuable tool with the potential to help them meet their overall science education goals, such as allowing for inquiry or more authentic learning (Lunetta, Hofstein, & Clough, 2007). However, Hofstein and Lunetta have shown that there seems to be very little evidence for this belief. There is a poor link between teaching laboratory activities and actual student learning (Hofstein and Lunetta, 1982). Traditional labs are often referred to as “cookbook” labs that ask students to follow the steps of a procedure without doing any design or analysis along the way. This contributes to the misunderstandings many students develop with regards to the study and practice of science. Hofstein and Lunetta suggest improving the level of inquiry in lab courses, which means allowing for more freedom for students to pursue scientific questions (2004).

Increasing inquiry (or the authenticity of the scientific experience) in the teaching lab can be difficult to achieve. One drawback of high-level inquiry activities is that they require a certain level of flexibility in both the instructor and the curriculum. To build this into the curriculum means having a wide variety of materials available for the many possible experiments a student...
might choose to conduct, which can be difficult to foresee as well as cost prohibitive, particularly in a large class. To ensure that an instructor has the background knowledge to aid students in their experimental design is another hurdle; in a middle school or high school setting, there may be difficulty in acquiring teachers with sufficient research experience to do this, and in a university setting (particularly with large classes), this is a high bar for first-time student instructors to clear. It also requires some level of monitoring to ensure similar experiences for all the students in different sections of large laboratory courses. Beyond that, the students in the class require a certain level of content knowledge for them to be able to envision an experiment to test a certain hypothesis. Ideally, there needs to be a small student-to-teacher ratio to allow for the level of guidance this requires. That is simply not possible in many schools. Therefore, the highest levels of inquiry are nearly impossible to achieve in an introductory-level teaching lab. If low-level inquiry lab activities are the only option, then, it is possible that explicit instruction on authentic science, including instruction on the nature of science, could help bridge the gap.

The Nature of Science

Understanding the nature of science is a difficult prospect and one that can span the work of an entire career. It is pursued by many different groups of thinkers: philosophers of science, historians of science, sociologists of science, science education specialists, and even practitioners of science. The perspectives of these groups can vary widely, and that leads to differing interpretations of what the nature of science actually is. One major interpretation of the nature of science insists that it refers to science as a way of knowing (Lederman, 1992). This leads to an emphasis on the epistemology of science and what can impact its development. “…scientific knowledge is tentative; empirical; theory-laden; partly the product of human inference, imagination, and creativity; and socially and culturally embedded,” (Lederman, 2002, p. 3).

Lederman and his many associates have developed a set of aspects of the nature of science that they feel is representative of the most important features of science as a way of knowing: empirical nature of scientific knowledge, inference and theoretical entities in science, nature of scientific theories, scientific theories versus laws, creativity in science, subjectivity in science, and social and cultural influences. Some in the field refer to these as the “Lederman Seven,” (Matthews, 2012). Lederman also routinely emphasizes that you cannot overlap the nature of science with science processes (2002). It refers only to “the values and assumptions inherent to the development of scientific knowledge,” (Lederman, 1992).

Lederman maintains a particularly staunch position on the nature of science that relies on the work of historians, sociologists, and philosophers of science. Abd-El-Khalick, a supporter of Lederman’s views, argues that the perspectives of scientists can be used to influence a definition of the nature of science, but they should not be privileged over the work of philosophers or historians of science (2012). That would lead to an overemphasis on the sociologist’s perspective and perhaps a lack of consensus on an established definition, as the views of each scientist would necessarily be colored by their specific experiences (Schwartz and Lederman, 2008). Abd-El-Khalick sees the input of scientists as only one source of information in crafting a nature of science definition.

There are many in the community who are somewhat critical of the Lederman definition, although it enjoys wide favor in the educational community for its ease of implementation and already-developed assessment materials. Some have objections from a philosopher’s
perspective. Matthews argues that these seven aspects chosen by Lederman are not particularly relevant to scientists and, furthermore, are not aligned purely in the epistemological realm (2012). There is obvious overlap of epistemological and sociological viewpoints here, and that makes philosophers of science uneasy. He further argues that the nature of science cannot be boiled down to a strict list of attributes; its nature varies from case to case, from domain to domain. His solution is to think of science as having various features that are much more fluid in terms of application to different situations. This allows educators, particularly, to focus on what is most important about science in a particular instance. Rudolph goes further, suggesting that a consensus view of the nature of science is not possible, and students should be exposed to the various aspects of science particular to specific domains of study (2000).

There are also objections to the Lederman definition from a science practitioner’s perspective. A definition of the nature of science that includes the perspectives of scientists is given by Osborne et al. (2003) and Wong and Hodson (2009, 2010, 2014). It is comprised of three areas: the nature of scientific knowledge, scientific inquiry (or scientific practice), and a view of science as a social practice. (Recall that the Lederman definition ostensibly only included the epistemology of science.) These researchers approach science as something being actively practiced; therefore, the perspective of current scientists is crucial to fully defining the nature of science. To elaborate, the nature of scientific knowledge involves a number of ideas, including the construction and use of theories and models, the certainty and uncertainty of scientific knowledge, the specific language used by the scientific community, the historic nature of the construction of scientific knowledge, and the role of creativity in that construction, among others. Scientific inquiry is comprised of the methods of scientific investigation utilized by the scientific community (or more specifically, the various discipline-specific communities within the larger scientific community). The roles of hypothesis and experiment are important components of scientific inquiry, as is the role of observation, the use of inference in data interpretation, the relationship between science and technology, and consideration of the human factor in conceiving of and performing experiments. Viewing science as a social practice involves considering the impact of society on science (such as how certain areas of study are chosen by scientists and funded by public and private groups and how the public’s moral and ethical positions can have major implications for which experiments are conducted) as well as the impact of science on society (such as the products of science and technology as well as the environmental effects of scientific efforts). It is also important to remember that science is a largely collaborative effort, not only in the way it is performed on small (in research groups) and large (as part of a scientific community) scales but also in the way its knowledge base is compiled. An understanding of even just a few of these ideas can have a major impact on a students’ personal relationship with science, on the level of a student in a lab course as well as a citizen in a global community impacted by the scientific community’s actions.

In reviewing the many viewpoints on what the appropriate definition is for the nature of science, it is clear that there is definite overlap in these researchers’ perspectives. There are more extreme views (i.e. developing a consensus view is necessary versus developing a consensus view is impossible), and there are views that emphasize various sources for the definition (i.e. historians/philosophers/sociologists of science versus active scientists). Between the Lederman and Hodson/Wong/Osborne definitions, there are some aspects of the nature of science that are found in each. In fact, they say many of the same things from differing perspectives. For the
development of a teaching laboratory curriculum, I find that the views of practicing scientists are invaluable, particularly in instructing potential future scientists. I also contend that because science is an eminently practical tool (as it really does no good if it is not used), deemphasizing the practice of science in defining the nature of science is detrimental to forming a complete understanding of science. Using the practice of science as a metric for how to define the nature of science makes it easier to point out to laboratory students how the nature of science can impact their work. It emphasizes the importance of learning about science on a philosophical level, and that may perhaps lead to more thoughtful and ethical scientists in the future.

Situated Learning

The curriculum assessed in this dissertation has been crafted in the spirit of situated learning (Lave & Wenger, 1991). This perspective asserts that learning cannot be separated from the context in which it occurs; in fact, the context plays a necessary part in learning. Knowledge is only useful in that it can be applied in particular situations. Therefore, context is not something to be scrubbed from the knowledge accrued during a learning episode; it is an inextricable part of shaping the knowledge that comes to be held by a person, and it shapes how that knowledge will be applied in future situations. This is particularly important to remember when designing a curriculum in a practice-based course, such as a chemistry laboratory. Much of the knowledge gained in this type of course is only useful in that it could be applied by students in future scientific situations (such as a science course, a research lab, or personal research on everyday scientific topics). Therefore, we as designers must account for the fact that students are learning about these topics in the context of a general chemistry laboratory course but need to find meaning in how these topics could be applied to their future lives. By framing the course in terms of NOS, we are encouraging students to recognize the higher value of these lab activities by relating them to overarching ideas about science that are transferrable to other disciplines and situations.

The Nature of Science in Education

The NOS literature has some clear findings in terms of constructing a curriculum that addresses NOS issues. To lead to actual gains in a student’s understanding of the nature of science, NOS instruction must be explicit; NOS topics cannot be learned simply by engaging in inquiry activities (Khishfe & Abd-El-Khalick, 2002; Schwartz, Lederman, & Crawford, 2004). This requires unpacking. Much of the NOS education literature involves studies that examine the impact of inquiry activities on students’ understanding of the nature of science. However, simply directing students to engage in inquiry is not enough for them to learn about the nature and generation of scientific knowledge. These topics must be covered explicitly in the curriculum for students to show gains in understanding in these areas. What does it mean, though, to say NOS instruction must be explicit to be effective? This is a well-established assertion of the NOS literature (Eastwood et al., 2012). Schwartz, Lederman, and Crawford identify explicit inquiry-based pedagogical approaches as paying “specific attention to NOS aspects,” (2004). They mention that implicit messages about NOS are present in almost any science lesson, but without calling actual attention to it, very little gain in student understanding of NOS material is achieved. Khishfe and Abd-El-Khalick make similar assertions, noting that explicit instruction “does not refer to didactic or explicit teaching strategies, but is meant to highlight the notion that NOS
understandings are cognitive instructional outcomes that should be intentionally targeted and planned for” (2002). Thus, explicit instruction does not refer simply to lectures on the subject from the instructor. As is the case with almost all other subject matter, some level of engagement with the topics is required for deep understanding to occur.

Contextualization of NOS instruction is a matter of debate. Studies suggest that NOS instruction can be effective either as direct instruction or as part of the context of the lesson (Eastwood et al., 2012). However, Eastwood et al. also notes that the NOS understanding gained by students in the contextualized version of the curriculum in their study was more grounded and applicable than that gained by students in the abstracted version. Clough has developed a scale for the effectiveness of the level of contextualization of NOS instruction, suggesting that perhaps a combination of techniques, from abstracted, informative instruction to fully embedded instruction, is most useful (2006). Combining these findings with the situated learning framework mentioned earlier in this chapter, the importance of crafting this curriculum to include embedded instruction on NOS topics becomes clear.

The literature is also very clear that opportunities for reflection on NOS instruction is a crucial part of student learning gains in the addressed topics (Schwartz, Lederman, & Crawford, 2004). This can take many forms, from discussion to application of NOS topics (Khishfe, 2013; Khishfe & Abd-El-Khalick, 2002; Yacoubian & BouJaoude, 2010). A last note on curriculum design requires the acknowledgement of the role of instructors in the effectiveness of NOS instruction. The nature of the connection between instructor beliefs and student learning gains in NOS is unclear. The instructor’s beliefs do not always manifest in his or her instruction, as demonstrated by Southerland, Gess-Newsome, and Johnston (2003), but the fact that there is a connection is apparent (Abd-El-Khalick & Lederman, 2000). Overall, the literature suggests that an explicit, embedded, and reflective NOS curriculum is the most likely to lead to improvements in students’ understanding of NOS topics.

Finally, while NOS instruction is suggested to help improve student attitudes, motivations, and content learning (i.e. Driver, Leach, Millar, & Scott, 1996), not many of these topics seem to be pursued in the literature. The findings of these studies are almost all focused on the development of NOS curricular materials and gains in NOS understanding. These gains are also generally measured in a decontextualized way without asking students to connect their understanding of NOS to scientific practice. What other impacts could NOS instruction have on students? Specifically, how might it impact their real-time experiences in the laboratory?

The Nature of Science in the Laboratory

There is very little work that examines the impact of NOS instruction specifically in a laboratory environment. Studies have shown that students’ epistemological beliefs can have noticeable impacts on their lab practices (Havdala & Ashkenazi, 2007; Vhurumuku, Holtman, Mikalsen, & Kolsto, 2006). This suggests that a student’s understanding of the nature of science influences his or her practice of science. More broadly, as mentioned above, the literature on laboratory courses suggests there is a need for more authentic scientific experiences for students (Singer, Hilton, & Schwingruber, 2005). Including explicit nature of science instruction in a laboratory curriculum could address both of these issues. It could impact a student’s epistemological beliefs, perhaps improving his or her practice in the lab, while making it easier for students to connect their lab experiences with authentic science.
There is evidence that NOS instruction in an inquiry setting is particularly helpful in improving students’ understanding of NOS (Schwartz, Lederman, & Crawford, 2004; Yacoubian & BouJaoude, 2010). It could help to make the NOS topics more accessible (Khishfe & Abd-El-Khalick, 2002). The more closely a student’s experience aligns with an authentic scientific experience, the more likely he or she is to improve his or her understanding of NOS (Akerson & Hanuscin, 2007; Russell & Weaver, 2011). Practicing scientists are the most likely to fully grasp NOS concepts such as the value of unexpected experimental results and the reality of scientific practice as model building (Samarapungavan, Westby, & Bodner, 2006). Thus, including NOS instruction in a high-inquiry lab activity could be extremely beneficial for students.

However, it is important to note that it is not possible for all laboratory courses to have many high-inquiry activities. Short of following an apprenticeship model or placing all lab students in an actual research laboratory, it is not feasible to give all introductory science students a truly authentic lab experience. There are both content-based and resource-based restrictions. However, an explicit and contextualized discussion of the nature of science might account for those missing pieces of a full inquiry experience. By discussing it outright and helping students see how their actual lab activities fit into the larger picture as seen by a practicing research scientist, students may receive a better understanding of the nature of science as well as a deeper appreciation for the laboratory course and how it supports more traditional science instruction. In a sense, including NOS instruction might impact the authenticity of a student’s experience in lab.

From this review a few key ideas about building a laboratory curriculum including NOS instruction become apparent. First, curricula that aim to improve students’ understanding of NOS topics must be explicit and reflective. Evidence from various analyses coupled with a situated learning framework promote embedding these NOS topics in the context of the laboratory. Supported by Matthews (2012), selection of the most applicable NOS topics as a focus for the curriculum is an acceptable method for crafting the materials of the course. Finally, attention to the implementation of such a curriculum is necessary to ensure the best possible student outcomes. The design of the curriculum analyzed in this study is the focus of the next chapter.
Chapter 3. Curriculum Development and Assessment

Figure 1. Timeline of the general chemistry laboratory renovation and curriculum development project at UC Berkeley.

History

For the past several years, our research team at UC Berkeley has been working to update the curricula for our general chemistry laboratory courses. These include Chem 1AL (general chemistry lab for non-chemistry majors), Chem 1B (the second semester of general chemistry lab for non-chemistry majors), and Chem 4A and 4B (the two semesters of general chemistry lab for majors). This process began with a grant from Cal-EPA and a desire to make the general chemistry labs greener. A few years later, we received funding from Dow to renovate the general chemistry lab space and also to overhaul the curriculum. The original goals of this overhaul were to increase attention to green chemistry in the experiments as well as to give students more opportunities for experimental design.

The curriculum design process has been iterative and continues even now. It began with developing ideas for new or revised experiments by looking to the literature. Teams of undergraduate students, led by graduate student researchers, helped test these ideas to create procedures for the new experiments. These teams also contributed to developing the instructional materials that would be used in these general chemistry courses. The new labs were then conducted in the courses, and feedback on these experiences was collected through conversations with the instructors and GSIs as well as survey responses from students taking the course. Our team would subsequently update the experiments and materials to address any difficulties encountered, and the second version of the labs would be tested the following semester. Essentially, this project has followed the principles of design-based research (Sandoval, 2014).
Course Structure

The focus of the studies presented in this dissertation is Chem 1AL. This is the general chemistry laboratory course for non-chemistry majors, and it is offered in the fall, spring, and summer sessions. The fall session is usually larger, with approximately 1200 students enrolling. The spring session is smaller, servicing around 700 students. The majority of students in these semesters identifies as a life science major, though there are a significant number of engineering, natural or environmental science, and health policy majors. Because the course is so large, it relies on an instructor and several GSIs (graduate student instructors) to implement the curriculum. The instructor teaches an hour-long lab lecture that all students must attend once a week. The students are also grouped into lab sections of 25-30 students that are taught by a GSI. There are often Teacher Scholars—undergraduates who have taken the course previously and can act as a peer mentor to the students and an aid to the GSI—assigned to a section, as well. The lab sessions last three hours, and students attend these once a week.

Before each lab session, a pre-lab exercise must be completed online by the students to ensure they have read through the materials prior to coming to lab. This helps them become familiar with the chemistry content, lab practices, and techniques to be employed in that session. It also makes them aware of waste procedures and safety concerns pertinent to that week’s experiment. The students receive a short lab lecture from their GSIs and, after completing the lab, turn in notebook pages to allow the GSIs to track their participation in the lab portion of the course. Students then turn in report sheets, which contain questions to help students analyze the data they have collected in that week’s lab and expand upon the topics they encountered by asking them to consider them in a different context. The hope is to help students achieve deeper learning on these topics. Depending on the semester, there can also be additional assignments, such as formal reports or argumentation exercises, scattered throughout the semester. Finally, each semester has a final lab exam conducted at the end of the semester that focuses on content, experiments, and techniques students used during the course.

The instructor varies each semester and has the ultimate decision on the curricular materials used. She conducts weekly GSI meetings to discuss important points in the experiments and how she would like the GSIs to broach the major themes of the course. Therefore, the instructor plays a major role in choosing and implementing the curriculum for Chem 1AL each semester. Similarly, the GSIs are the experts seen most often by the students in the course. They are the implementers of the curriculum that are closest to the students, meaning their decisions on what topics to discuss and how to discuss them in the lab are a significant part of the curriculum implementation chain.

Redesign Goals

The major goals of the curriculum redesign were to increase the green chemistry portion of both the design of the experiments and the learning goals of the course along with opportunities for students to engage in experimental design. The green chemistry goals are currently being assessed by other members of our research team. The focus of the studies that follow is the experimental design portion of the redesign and how we went about designing and assessing these experiences for the students.

Experimental design activities appear specifically in two labs. During these sessions students must design their own experiments in order to answer a chemistry-based research
question. One of the labs (Polymers 2—see Appendix 1 for more details) asks students to design a toy for children where they must optimize a particular polymer property, such as bounciness or stickiness. The other lab (Extraction of Curcumin) requires students to explore the extraction of curcumin from turmeric using various solvents and methodological choices. After a few rounds of implementation of these labs, the research team decided to work on scaffolding them to help students with the challenges of designing their own experiments. It is a difficult activity for students in an introductory chemistry class, as their understanding of chemistry content and experimental techniques is limited. Their knowledge of experimental design itself—what variables to choose, how to control for everything but one variable, how to analyze data to answer a research question—is also not extensive, as survey data shows most of them have had no experience with this type of high-inquiry activity before this course. Therefore, we chose to include explicit instruction on the nature of science to address these problems, with me as the primary designer.

Including NOS instruction in the course would hopefully help students with these high-inquiry labs by allowing them to learn more explicitly about topics of scientific knowledge and inquiry and how they relate to each other. Perhaps if students understood the purpose of experiments—to test scientific models and lead to an addition or refinement of scientific knowledge—they would have a better sense of what choices they should make for their variables or of the limitations of the interpretations they could make from their data. These topics are often considered desirable as learning goals in and of themselves for lab courses. Thus, there was motivation to include such instruction for that reason, as well.

**Development of NOS Materials**

As noted in the previous chapter, for students to make significant learning gains in the area of NOS, the instruction (and materials) must be explicit, embedded, and reflective. Below is a description of how I approached these three tasks:

**Explicit:** I began by addressing NOS in an introductory page in the manual. Here, the three parts of NOS were defined: the nature of scientific knowledge, the nature of scientific inquiry, and the nature of science as a social practice. Essentially, this provided students with a definition of the nature of science, and it was referenced in every lab in the course. Additionally, the same structure was used to describe each lab in the manual:

- research question
- learning goals
- green chemistry notes
- NOS notes
- chemistry content background
- experimental procedure
- report sheet

All the labs follow this same set up (see Appendix 1 for the full text of two of these labs) in the hopes that students would benefit from seeing the general concepts of scientific practice week
after week: start with a research question, read about the topic, collect data with experiments, and analyze that data to answer the research question.

In the NOS notes at the beginning of each lab, which of the three parts of NOS that were being addressed in that lab were identified specifically, along with how they related to the work the students were about to do. Thus, the NOS topics were being explicitly explained in the context of the lab. Additionally, instructors and GSIs were encouraged to discuss NOS topics specifically in their interactions with the class. Students therefore received instruction via both the materials and the instructor, depending on who taught the course that semester.

*Embedded:* As was mentioned in the section above, NOS topics were explicitly linked to the lab activities being conducted by the students. Students were also asked to answer questions related to these NOS topics using either the data they had collected from the lab or the experiences they had acquired from performing it. For example, in the Polymers 1 lab students were asked to write a scientific model for the experiment they were performing to determine the relationship between stretchiness and the amounts of PVA and guar gum in a solution (see Appendix 1). In this way NOS was embedded into the context of the labs.

*Reflective:* Finally, opportunities for students to reflect on NOS topics were built into the curriculum in two ways. First, students were asked to engage with NOS topics throughout the lab activity. For example, in Polymers 2 students had to develop, test, and refine scientific models as part of the activity itself, rather than requiring them to develop experiments solely. Second, many of the questions in the report sheet required students to reflect explicitly on NOS topics. Once again referencing Polymers 2, students were asked to reflect on the social aspects of science by considering their own experiences reporting their procedures to others and agreeing on the parameters for their polymer property tests as a group.

**Summary of Labs with NOS Instruction**

Below is a summary of the labs developed for the new Chem 1AL curriculum containing explicit NOS instruction.

- **How the Nose Knows:** Students collect data as a class, guided by the GSI, to determine a model for how the sense of smell works on a molecular level. This is done iteratively, with the class developing a model, testing it, and refining it several times. Students are introduced to scientific models and the concept of model building in the introductory materials and asked to engage with them during the course of the experiment.

- **Polymer Properties and Applications Part 1:** In the first part of this lab, students work in pairs to qualitatively assess various polymer solutions for several properties, including stretchiness, bounciness, stickiness, and viscosity. They can track these property changes with additions of various compounds and mechanical variables, such as stirring. This provides a base of content knowledge for the following week. The second part of this lab is an experiment asking students to mix a range of PVA to guar gum ratios with borax and to determine how this ratio impacts stretchiness. While everyone completes the same experiment and is provided the ratios and procedure, they must all develop an initial
model for what this experiment is testing, thus engaging with NOS topics in the context of the lab.

• **Polymer Properties and Applications Part 2**: This lab builds on the lab prior by asking students to use the trends they observed to build a toy that has optimized a particular polymer property. This is a high-inquiry activity that builds on the previous two labs. It requires students to engage in the same model building process as they did in “How the Nose Knows” with less guidance while also developing initial and final models for their properties, as they did in “Polymers 1.” Here, students are heavily engaged with NOS topics to execute the lab activities.

• **Biofuels Part 1: Ecotoxicity Assay Setup and Synthesis of Biodiesel**: This is the beginning of a three-part lab examining the topic of biofuels. Students collect data throughout the series to make an argument about biofuels in a future exercise. The first week explores topics of toxicity and synthesis. They set up an ecotoxicity assay to test the toxicity of several biofuels and also begin a biodiesel synthesis. They engage with NOS topics around science as a social practice by considering how science influences society and vice versa, given the publicity of the topic of biofuels.

• **Biofuels Part 2: Isolation of Biofuels and Analysis of Ecotoxicity Assay**: In the second lab of the series, students analyze the data from their assays and complete the synthesis of their biodiesel samples. They reflect on the many steps that can be required for a synthesis and how important toxicity considerations are for a fuel.

• **Biofuels Part 3: Calorimetry**: In the final lab of the biofuels arc, students burn their biodiesel samples, along with a few other biofuels, to determine fuel efficiency via calorimetry. They also test the viscosity of the fuels, both as a consideration for the ease of use of the fuel as well as a test of their synthesis skills. They begin to consider how to use all this data they have collected to make an argument for the best biofuel.

• **Biofuels Argumentation**: This is an additional assignment that students complete after collecting all their data from the biofuels unit. Students engage with argumentation as they reconcile sometimes disparate data to make an evidence-based argument in favor of a particular fuel.

• **Acids in the Environment Part 1: Preparation and Properties of Gases**: This is another three-part series of labs examining acid-base chemistry in the context of the environment. In part 1 students examine Henry’s Law and the solubility of carbon dioxide in water in an effort to understand the acidification of bodies of water. They are introduced to the ideas of pH and equilibrium. They also must think about different methods of collecting and interpreting data, as these can change the conclusions a scientist makes.

• **Acids in the Environment Part 2: Determining Molarity**: In part 2 students execute an indicator titration to model the destruction of coral by ocean acidification. It heavily explores acid-base ideas and the calculations involved in acid-base reactions. Students are also asked to consider sources of error and the impact experimental methodology can have on the data collected.

• **Acids in the Environment Part 3: Potentiometric Titration**: The final lab of this sequence asks students to perform potentiometric titrations of various acids related to ocean acidification. They are introduced to titration curves and collect class data to compare
characteristics of different acids. They continue their consideration of methodology choice and how a scientist would go about making the right choice for his or her research question.

- **Extraction of Curcumin from Turmeric and Spectroscopic Analysis**: The final lab of the course comes full circle, asking students to design an experiment to maximize the extraction of curcumin from turmeric. They explore Beer’s Law, serial dilutions, and spectroscopic techniques to determine the concentration of the solutions they are making. They are asked to choose a research question, create an initial model for extraction, and conduct experiments of their own design before gathering data and analyzing it to make a final conclusion about the chemistry behind extraction.

Below is a table showing the areas of NOS covered in each of these labs.

<table>
<thead>
<tr>
<th>NOS Topic</th>
<th>HTN</th>
<th>P1</th>
<th>P2</th>
<th>Bio 1</th>
<th>Bio 2</th>
<th>Bio 3</th>
<th>Acids 1</th>
<th>Acids 2</th>
<th>Acids 3</th>
<th>Extraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Inquiry</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Social Practice</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

*Table 1. Breakdown of NOS topics in redesigned Chem 1AL labs.*

**Assessment**

Once the materials for Chem 1AL were developed, the research team wanted to assess several student outcomes.

- Learning gains: chemistry content, lab techniques, green chemistry knowledge, NOS understanding and application
- Reception of curriculum: enjoyment of labs, favorite/least favorite experiments, most valued outcome, etc.
- Student experience: use of resources, valued resources, study methods, confidence in doing chemistry and performing or designing labs

Several types of data were collected to attempt to assess the curriculum. These included survey responses, student work, interviews of students, and video and audio of GSIs teaching in the lab. While much of this assessment work is still being done, I assessed the design and implementation of the NOS portion of the curriculum. As Table 1 shows, many NOS topics were covered over the course of ten labs. Analyzing all the NOS topics from the course and all the methods of implementation was not possible for a single dissertation. Therefore, I chose to focus on analyzing two particular areas: scientific models and model building.

In this course scientific models are presented in the context of NOS. Their creation and refinement are among the most important goals of practicing scientists. Various aspects of models are addressed throughout the course, including their nature, their purpose as explanations of the natural world, and their use in scientific practice to drive experimental design during investigations. These ideas align well with the literature (Krell, Upmeier zu Belzen, &
The course materials also represented scientific models as molecular-level explanations of macroscopic phenomenon. This is certainly not true of all scientific models, but it was an appropriate definition for a general chemistry course. It also pushed students to consider the chemistry behind macroscopic observations they were making in the lab.

Model building refers to the process of creating and refining a scientific model through the construction of hypotheses and experiments. Addressing this topic was a way to scaffold the experimental design activities students had to complete in Polymers 2 and the Extraction lab. The materials were designed to aid students in their experimental designs by starting with an initial model that they could use to choose variables around which to craft experiments. They finished by updating their original models to create a final model.

The choice to focus on student’s understanding of models and model building was made for two main reasons. First, these are topics covered in some depth over the course of the semester. They play a particularly important role in three labs early in the course: How the Nose Knows, Polymers 1, and Polymers 2. These labs build upon each other, each discussing models and model building in increasing detail until students are asked to create their own models and test them with experiments of their own design. Therefore, these subject areas received significant attention from the students. Second, scientific models and model building are a major part of experimental design. Determining students’ understanding of these topics is helpful for understanding how this instruction impacted their scientific practices, a major goal of the curriculum redesign. The following research questions were inspired by the curriculum development and assessment process around the inclusion of the NOS topics of scientific models and model building.

Research Questions

*Did this curriculum result in learning gains for students on the topic of scientific models? How did students respond to this curriculum?*

*How well are students able to write scientific models? How important were the instructional choices of the lab instructors and GSIs in impacting students’ abilities to do this?*

*How did this curriculum impact students’ experimental design approaches in a high-inquiry laboratory experience? Specifically, did the requirement for students to write an initial and final model before and after conducting their own experiments lead to logical, thoughtful choices in their experimental design?*

These questions are answered over the course of the next three chapters. Each one required different data sources and methods of analysis, so these topics are detailed separately for each question in its corresponding chapter.
Chapter 4. The Positive Impacts of Explicit Instruction on Scientific Models in a Laboratory Curriculum

Introduction
Scientific models and model building are important elements of scientific practice. Scientific models are explanations of phenomena and represent an important portion of scientific knowledge. Models frame scientists’ hypotheses, experimental designs, and data analysis choices, leading to new conclusions that reshape their initial models and add to the body of scientific knowledge. All of this is done within a social and cultural network that helps to determine the validity and placement of this new knowledge within existing knowledge bodies (Krell, Upmeier zu Belzen, & Kruger, 2014; Treagust, Chittleborough, & Mamiala, 2002).

The topic of scientific models is not usually included in laboratory courses. There are simply too many other topics that are considered vital for students to learn. Thus, instruction on models is not the norm in general chemistry laboratory courses. This research investigates the value of adding models to a laboratory course for student learning. It attempts to determine the effectiveness of a curriculum that includes explicit instruction on scientific models and model building. This curriculum was used for one semester in Fall 2017, so the results of this analysis will be specific to that group of students and that instructor’s particular implementation choices.

Research Question
This research assessed students’ learning gains around scientific models, as well as their reactions to the materials themselves.

Did this curriculum result in learning gains for students on the topic of scientific models?
How did students respond to this curriculum?

These questions were investigated by analyzing answers to pre- and post-course survey questions given to the students of Chem 1AL (general chemistry laboratory course for non-chemistry majors) at UC Berkeley during the Fall 2017 semester. They probe students’ understanding of scientific models and their evaluation of the course as a whole.

Curriculum Materials
Scientific (more specifically, chemical) models are referred to in this course as molecular level explanations of macroscopic phenomena. This is an area that is covered multiple times in the course, specifically in the How the Nose Knows, Polymers 1, and Polymers 2 experiments. A deep understanding of scientific models includes an understanding of the nature of scientific knowledge, the nature of scientific inquiry, and the meeting of these two areas. The curriculum explores modeling beginning with How the Nose Knows, which introduces students to the idea of scientific models and allows them to engage in model building in a highly guided fashion (see Chapter 3 for a description of all labs). The graduate student instructor (GSI) teaching that lab section helps students to develop an initial model, design experiments to test hypotheses predicted by the model, and ultimately reach a more detailed and evidence-based model about how the sense of smell works on the molecular level. The Polymers 1 lab (Polymer Properties and
Applications: Part 1) expands on these ideas by introducing students explicitly to the process of model building and asking them to write a model for a prescribed experiment they all must complete concerning a macroscopic polymer property. Finally, the Polymers 2 lab (Polymer Properties and Applications: Part 2) is a continuation of Polymers 1 that asks students to design their own experiments to test an initial model of their own devising about a polymer property. They must develop their own hypotheses and experiments, conduct their investigations, analyze their data, and make conclusions that ultimately lead to a final, refined chemical model for their polymer property.

Modeling is discussed extensively in three experiments, and this arc of activities ends with a high-level inquiry exercise in experimental design that requires students to create and refine a chemical model. Therefore, it is safe to conclude that students have had a significant amount of exposure to the idea of scientific models. However, they are also exposed to many other important topics, including chemistry content, laboratory skills, and others that are considered necessary learning goals for a laboratory course. Because of that, it is possible that students may not make particularly deep learning gains related to scientific models.

Methods

Participants:

Chem 1AL is the laboratory course for general chemistry for nonmajors. The majority of the students in this course identify as life sciences/biology majors, with a significant number of engineering majors and a smaller number of public health, social science, and environmental and natural science majors. The exact make-up differs between the fall and spring semesters, particularly with regards to the engineering population. The fall semester typically has a larger enrollment, with around 1200 students enrolling in the course. The spring semester enrollment is usually around 700 students.

![Chem 1AL Major Designations](image-url)
Instructor:
The course consists of one overall instructor who conducts a one-hour lab lecture that all students must attend once a week. The students also attend a three-hour lab section once a week that is taught by a GSI. Each section is approximately 25-30 students, and in Fall 2017 there were 40 sections with 20 different GSIs. Almost all sections also have a Teacher Scholar, an undergraduate student who has taken Chem 1AL before and can act as a peer mentor to the students and an aide to the GSI, who is often new to teaching this course.

Student Consent:
As prescribed by the IRB for this project, in Fall 2017 consent was obtained from students to participate in this study through a survey that was distributed electronically through the course website at the beginning and end of the semester. Only students who consented to participate in the study and to allow the research team to utilize their work from the course were included in the analyses presented here. This resulted in approximately 550 consenting students of the 1200 enrolled. The analyses that follow represent populations that range between 445 and 520 students, depending on how many students answered those particular questions on the survey. Unfortunately, there was no way to obtain this demographic data for the entire course population.

Prior Participant Experience:
The figure above shows the distribution of students’ past experience with chemistry. Students were categorized as having less than 2 semesters of chemistry, 2 semesters of chemistry, or more than 2 semesters. The data shows that the vast majority of students had completed at least two semesters of chemistry in their past. Very few were completely new to chemistry.

Figure 2. Past chemistry experience of Fall 2017 Chem 1AL students.
It is clear that the majority of students in Fall 2017 had completed a chemistry lab prior to Chem 1AL. Most students had only followed a written procedure rather than designing their own. Therefore, these students had some experience with which to form expectations of a lab course, but over half of them did not have much experience with authentic scientific practice.

Data Collection:
Surveys were administered at the beginning and end of each semester with an offer of bonus points in the course upon completion. They were used to gather demographic information, details of students’ past experiences with chemistry, and answers to course evaluation questions. This is also where students were given the opportunity to consent to participating in the study.

Data Analysis: Scientific Models
For this study data from the pre- and post-course surveys of the Fall 2017 cohort of Chem 1AL was analyzed. The first goal of this analysis was to determine if students made learning gains on the topic of scientific models. Models are discussed extensively in three experiments, and this arc of activities ends with a high-level inquiry exercise in experimental design that requires students to create and refine a chemical model. Therefore, it is safe to conclude that students have had a significant amount of exposure to the idea of scientific models.

The same question was asked in the pre- and post-survey about scientific models: “In your own words, what is a scientific model?” Instead of developing a coding schema from scratch, I chose to analyze the data using a definition of scientific models from the literature. This was done for two reasons: first, to compare students’ answers to an accepted idea of scientific models in the science education community (rather than to the research team’s own, internal definition of scientific models). This helped to validate the findings, as this schema was validated in a previous study (Krell, Upmeier zu Belzen, and Krüger, 2014). It was also done to ensure there was a comprehensive definition with which to measure students’ ideas about models. Many different aspects of scientific models are touched on in the curriculum. While it was unlikely students would cover all these ideas in their short survey answers, it was important to be able to document their understanding of whichever areas they chose to mention.
<table>
<thead>
<tr>
<th></th>
<th>Descriptive</th>
<th>Predictive</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Level 1</td>
<td>Level 2</td>
</tr>
<tr>
<td></td>
<td>Score 1</td>
<td>Score 2</td>
</tr>
<tr>
<td>Nature of models</td>
<td>Replication of the original</td>
<td>Idealized representation of the original</td>
</tr>
<tr>
<td>Multiple models</td>
<td>Different model objects</td>
<td>Different foci on the original</td>
</tr>
<tr>
<td>Purpose of models</td>
<td>Describing the original</td>
<td>Explaining the original</td>
</tr>
<tr>
<td>Testing models</td>
<td>Testing the model object</td>
<td>Compare the model and the original</td>
</tr>
<tr>
<td>Changing models</td>
<td>Correcting the defects of model object</td>
<td>Revise due to new insights</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Score</td>
<td></td>
</tr>
<tr>
<td>Molecular nature of</td>
<td>No mention of molecular nature of chemical models</td>
<td>Explicit recognition of chemical models as</td>
</tr>
<tr>
<td>chemical models</td>
<td></td>
<td>molecular</td>
</tr>
</tbody>
</table>

Table 1. The “model of model competence,” originated by Upmeier zu Belzen and Kruger (2010) and reproduced by Krell, Upmeier zu Belzen, and Krüger (2014). The score row indicates how students’ answers that fell into these categories were scored. An answer that did not have any evidence of understanding of that aspect of scientific models received a score of 0 for that category. An answer that fell into the Level 1 category for a descriptive aspect received a score of 1, and an answer that fell into Level 2 for the descriptive aspect received a score of 2. The predictive score for all responses was a 0 for answers with no evidence of a predictive understanding of scientific models or a 2 for answers that did show such evidence. The molecular nature of chemical models is a category added specifically for this study to account for the explicit instruction in the curriculum that the most explanatory chemical models contain a molecular-level description of a macroscopic phenomenon. It has been coded on a binary scale.

The table above shows the coding schema for scientific models. Students answered the survey question in 1 to 2 sentences. Each category has a descriptive and predictive score. As explained by Krell, Upmeier zu Belzen, and Krüger, scientific models have both a descriptive and predictive nature (2014). Understanding the predictive side of models does not imply an understanding of the descriptive side, so it made sense to code for these concepts separately. The descriptive score is on a scale of 0 to 2, and the predictive score is binary—either a 0 or 1. However, the predictive score was weighted to be a 0 or 2 to equate the highest predictive score with the highest descriptive score. Essentially, a student could write an answer that indicates an understanding of the predictive nature of models but not the descriptive nature of them. Scoring that answer as a 1 would equate it with an answer that represents a lower-level understanding of the descriptive nature of models, when in actuality, understanding the predictive nature of models is more difficult and equivalent with the higher-level understanding of their descriptive nature. In essence, there is not a category for a lower-level understanding of models’ predictive nature, so the coding for predictive nature must be either a 0 or 2 to accurately represent a
student’s understanding of models. Therefore, a student who shows an understanding of both the descriptive and predictive nature of models could receive a score of 3 or 4.

I also added a category to code for chemical models specifically. One idea that was emphasized during the course was that scientific models in chemistry many times involve a microscopic or molecular level explanation of a phenomenon. This is not an idea that fits nicely in the categories of the predetermined schema, so a molecular level category was added. Students scored either a 0 or 1 for this as a way to indicate if they mentioned this topic or not. Answers were not coded on a 0 or 2 binary scale, as they were in the predictive category, since the molecular category is meant to represent a separate concept from those explained in the model of model competence. It has no part of the ordinal relationship between the descriptive and predictive coding categories, so a score of 1 for a student does not misrepresent his or her level of understanding. Interrater reliability tests were run for coding the purpose category of models as well as the molecular level nature of models. Over 10% of students’ answers were coded, and a comparison between two unique coders for these three coding events (purpose: descriptive, purpose: predictive, and molecular level) yielded kappa values above 0.8 in each case.

The resulting scores for students are an overall representation made by adding students’ scores for each category together. The scores for each category, including the molecular level category, were added together to make an overall sum for each student. These sums were compared in a paired t-test to show if significant gains from pre- to post-survey were achieved.

High scores were not expected for students overall. This is a very detailed schema for scientific models. Because this course had several learning goals, only one of which was an understanding of the nature of science, the topic of scientific models (along with many other NOS topics) could not be covered in depth. However, all of these ideas about scientific models were touched on in some capacity in the course (either in the materials or by particular instructors as they presented the material to their students), so having a more comprehensive schema for coding students’ answers to this question was important. This helped to ensure that the understanding of students regarding scientific models at any level or in any capacity could be captured.

Data Analysis: Curriculum Reception

The survey questions were also analyzed to determine students’ reception of the curriculum. There were 3 Likert style questions with 5 possible answers: not at all, a little, somewhat, a good deal, or a great deal. They are as follows:

Please indicate how much you agree with the following statements.

- I enjoyed doing labs.
- The laboratory manual was clear.
- The labs enhanced my ability to design my own procedures and experiments.

There was an additional question that had either “yes” or “no” as possible answers:

Did you find the introductory materials in the lab manual before each experiment useful?
Finally, one free response question was analyzed:

“What was the most valuable thing you gained from lab?”

This was a free response question, and students’ answers were generally 1 to 2 sentences long. Categories were determined to fit each answer and help clarify how valuable students found the NOS aspects of the curriculum to be. These included lab skills (i.e. techniques, exposure to working in a lab, lab safety, lab etiquette, writing lab reports, etc.), group work, NOS topics, chemistry content, application of knowledge, and student skills, among a few other sparsely populated categories. The NOS topics category included mentions of topics such as experimental design, data analysis, and scientific practice. These references to scientific practice were generalized in their wording, rather than explicitly referencing chemistry laboratory procedures or practices.

Results and Discussion

*Scientific Model Learning Outcomes:*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Observations</th>
<th>Mean</th>
<th>Standard Error</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-model scores</td>
<td>479</td>
<td>2.3883</td>
<td>0.0628</td>
<td>1.367</td>
</tr>
<tr>
<td>Post-model scores</td>
<td>479</td>
<td>2.6472</td>
<td>0.0600</td>
<td>1.3140</td>
</tr>
</tbody>
</table>

$t=-3.2113$, df=478  
$p=0.0007$  
Cohen’s $d=0.10$

Table 2. Results of a paired, two-tailed $t$-test comparing students’ pre- and post- answers to the survey question about the definition of a scientific model. This indicates students did make significant gains in their understanding of this NOS topic, but the small effect size indicates these gains were minimal.

The first goal of this analysis was to determine if students made learning gains on the topic of scientific models. It is important to note that the answers to these questions were collected from a long and extensive survey. The answers were short, and survey fatigue for the students was a real possibility. Coupling that restriction with the reality that the topic of scientific models was only a portion of the many topics confronting students in the course each week, the analysis of this survey question using this schema represents a particularly conservative view of students’ understanding of scientific models. It would be difficult to discuss all the areas of models indicated in Table 1 in less than a paragraph. Because the curriculum was developed separately from the schema, students’ answers were not guaranteed to fit nicely into the various indicated categories. Therefore, any positive changes in students’ answers from pre- to post-course survey analysis would be considered a measure of success for this curriculum.

Table 2 indicates that overall, according to the schema for a scientific model developed by Upmeier zu Belzen and Krüger and described by Krell, Upmeier zu Belzen, and Krüger (2014), the students did make gains in their understanding of scientific models. As can be seen in the figure below, the largest gains were observed in the molecular level category and the descriptive purpose category. Other categories, such as multiple, testing, and changing models had very few entries from students, either pre- or post-course. The predictive side of all categories was also
sparsely populated. The curriculum focused on teaching students about molecular models, and other NOS topics, in the context of the laboratory. While scientists use them in their research, students in the course used them in the first few experiments of the semester in order to explore and explain observed phenomena. They also used them in their own experimental designs in the culmination of these activities. Therefore, it makes sense that students’ responses were heavily weighted toward discussion of scientific models in terms of how they relate to scientific practice. Example answers from the students can be found in Table 3. One of the ideas about chemical models that is prevalent in the course is that they very often involve explanations of the molecular-level activity of a phenomenon. Hence, a category was added to represent this aspect of models that was specifically addressed in the course. Given these course specifics, it is therefore unsurprising that most of the answers offered by students would pertain to the purpose and molecular-level nature of models.

It is important to note that the gains were very small when taking into account all coding categories. Perhaps looking at categories that were most heavily addressed in the curriculum would help pinpoint the areas where students made the most gains.

<table>
<thead>
<tr>
<th></th>
<th>Descriptive</th>
<th>Predictive</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Score</strong></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td><strong>Nature of models</strong></td>
<td>A visual aid that helps people understand potentially complex scientific processes</td>
<td>A scientific model is a representation of an idea or a system.</td>
</tr>
<tr>
<td><strong>Purpose of models</strong></td>
<td>A scientific model is used to describe a scenario that must be investigated</td>
<td>An explanation at a molecular level of a correlation between two properties.</td>
</tr>
<tr>
<td><strong>Molecular nature</strong></td>
<td>An explanation of how the microscopic properties of molecules and their interactions affect their macroscopic properties.</td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Example answers to the question, “In your own words, what is a scientific model?” These are representative answers for the categories of nature of models, purpose of models, and molecular nature of chemical models and the corresponding scores. Note these answers may also have been coded with scores for other categories, as well.
Figure 4. Pre- and post- scores on model understanding for students from Fall 2017 Chem 1AL. Pre-scores are blue series, and post-scores are orange (i.e. Nature1 vs. Nature2). These results are specifically for the descriptive nature of models, the descriptive purpose of models, and the molecular level nature of models. See Table 1 for details on the scoring rubric.

One other important trend to note from the data is the decrease in understanding that appears to occur in the descriptive nature category. This may be more an artifact of matching a thorough coding scheme with brief answers than any true loss of understanding. Describing the nature of scientific models in two sentences is challenging; measuring their understanding of this idea from such a short entry is nearly impossible. That means their scores in this category hinged largely on vocabulary: many students used the word “representation” to describe the nature of a scientific model. This indicated they had a level 2 understanding, according to the scheme above. However, it is difficult to know what they meant by “representation” without further elaboration. They could have changed their word choice to “replication” or “simulation” for the post-survey, which would earn them a level 1 score. This would show a decrease in understanding quantitatively, but in reality, it may only be a change in word choice. If they did not mean “representation” to be an “idealized representation” of a phenomenon, they were perhaps given too high of a score on the pre-survey unintentionally.

Given this and the small gains made by students in categories that were less emphasized in the curriculum, the table below shows the results of the paired t-test solely for the sum of the
purpose and molecular level measures of students’ understanding of models, ignoring the other areas of this topic which were not covered in depth in the course or that could not be confidently coded with this schema.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Observations</th>
<th>Mean</th>
<th>Standard Error</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>sum_pm_1</td>
<td>479</td>
<td>0.72</td>
<td>0.04</td>
<td>0.91</td>
</tr>
<tr>
<td>sum_pm_2</td>
<td>479</td>
<td>1.38</td>
<td>0.05</td>
<td>1.01</td>
</tr>
</tbody>
</table>

T=-11.6, df=478  
p<0.00005  
Cohen’s d=0.60

Table 4. Results of a paired, two-tailed t-test comparing the pre- and post-survey scores of students on the question, “In your own words, what is a scientific model?” This is specifically testing the difference in pre- and post-scores for the sum of the purpose and molecular level measures of students’ understanding of models.

From this t-test, it is clear that, for the area of scientific models that was most heavily represented in the course materials and instruction, students made significant gains from before the course to after its conclusion. These gains are even larger than those shown in the prior t-test shown in Table 2, suggesting that the curriculum has been effective at improving their understanding of scientific models, as the areas it most focused on showed the best improvement.
Curriculum Reception:

Figure 5. Student reception of the curriculum and its associated materials. These are all responses to questions from the post-survey given in Chem 1AL from Fall 2017 to evaluate students’ reactions to the curriculum. The questions are described below. The y-axis represents the number of students. The range of students responding for each of these questions is 509 to 511.

To track students’ reception of the new curriculum, they were asked several evaluative questions about their experiences with the materials and their implementation in a post-course survey. The figure above represents the range of responses to the following questions:

*Please indicate how much you agree with the following statements.*
- I enjoyed doing labs.
- The laboratory manual was clear.
- The labs enhanced my ability to design my own procedures and experiments.

*Did you find the introductory materials in the lab manual before each experiment useful?*
Answers to the first three questions could be 1 of 5 answers ranging from “not at all” to “a great deal.” Answers to the final question were either “yes” or “no.” It is clear from the figure that overall, students enjoyed the labs, found the manual to be clear, and found the introduction of each experiment (where much of the information about scientific models was disseminated) to be useful. Students were also asked to rate how much they agree with the statement, “The laboratory manual was confusing.” This served as a check for how consistently students were answering the survey questions. After reverse coding students’ answers to this question, they were compared quantitatively with the answers to the positive version of the question. According to a paired, two-tailed t-test (see the table below), there was not a statistically significant difference in the means of students’ responses to either of these questions, providing validation of the analysis.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Observations</th>
<th>Mean</th>
<th>Standard Error</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manual_clear</td>
<td>508</td>
<td>2.62</td>
<td>0.04</td>
<td>0.84</td>
</tr>
<tr>
<td>Manual_confusing</td>
<td>508</td>
<td>2.69</td>
<td>0.04</td>
<td>0.97</td>
</tr>
</tbody>
</table>

\[ t= -1.60, \text{df}=507 \]
\[ p=0.055 \]

Table 5. Results of a paired, two-tailed t-test between the quantified answers of the “clear manual” and “confusing manual” questions discussed above from the Fall 2017 post-survey.

Finally, students answered the question, “The labs enhanced my ability to design my own procedures and experiments,” positively. This suggests students found the labs to be helpful in impacting their understanding of this part of scientific practice, which was an important goal of the curriculum redesign.
Figure 6. Results of the analysis of the survey question, “What was the most valuable thing you gained from lab?” The y-axis represents number of instances a full or partial statement related to that category was made by a student. There was a total of 553 “things of value” mentioned by 475 students.

In the post-course survey students were asked, “What was the most valuable thing you gained from lab?” The largest percentage of answers mentioned that the students most valued lab skills. The next most popular answer category was group work: students really valued learning how to work in cooperation with others. Most interestingly for this study, the NOS category was the third most popular. There were 19 separate mentions of experimental design along with 21 explicit mentions of models, largely in reference to learning how to develop models. The fact that these comments showing that students valued learning about NOS topics came with no specific prompting about NOS suggests that including these topics in the curriculum was appreciated by a significant number of students.
<table>
<thead>
<tr>
<th>Category</th>
<th>Example Answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lab skills</td>
<td>Different titration techniques (indicator + potentiometric)</td>
</tr>
<tr>
<td>Group work</td>
<td>Performing labs in groups helped me understand the importance of teamwork.</td>
</tr>
<tr>
<td>NOS</td>
<td>The expertise in lab procedures, scientific models and hypothesis, and in the operation of Chemistry equipment.</td>
</tr>
<tr>
<td>Chemistry content</td>
<td>A better introduction to functional groups that I think I will use in future chemistry classes.</td>
</tr>
<tr>
<td>Student skills</td>
<td>I made good relationships with my classmates and GSI and learned how to follow instructions and perform a task in a given amount of time.</td>
</tr>
<tr>
<td>Application of knowledge</td>
<td>The most valuable thing I gained from lab was real world applications of chemistry lecture topics.</td>
</tr>
</tbody>
</table>

Table 6. Example answers to the survey question, “What was the most valuable thing you gained from lab?” These are representative answers for the most popular coding categories. Note that these answers may also have been coded into other categories, as well.

Conclusions and Future Work

Students have responded well to this curriculum and its explicit instruction on the nature of science. They found the manual clear and useful, the experiments enjoyable, and the topics of model building and experimental design valuable to learn. These findings are suggestive of a positive lab experience for students studying chemistry with this curriculum and may be reinforced with further interviews or survey analysis. Additionally, despite the conservative nature of the measure used in this study, students showed gains in their understanding of the topic of scientific models. This suggests that students can make learning gains in this topic with some simple steps to redesign an existing curriculum (see Chapter 3 for more details on the design of this curriculum). However, it is important to note that these results were obtained for one semester, and the implementation of the curriculum plays a major role in its effectiveness and reception. It is possible that this curriculum would not result in the same learning gains if different implementation choices were made by the instructor. Analyzing answers to the same questions from Spring 2017, a different group of students with a different instructor but a very similar curriculum, would help clarify these findings. It is also important to point out that this data is only representative of students who voluntarily agreed to participate in the study. The large sample size is impressive, but as there was no way to control who agreed to the study, it should be noted that this sample may not be representative of the entire Chem 1AL population that semester.

Building on these results, it should be noted that there are other ways to show changes in students’ understanding of scientific models. One way is to examine how students applied this understanding in the lab setting as they completed their experiments and report sheets. This would help validate the findings from this chapter by examining students’ understanding of models in a different context. If there are similar gains, the results from this chapter will be strengthened. An analysis of this type is conducted in the next chapter when the models students wrote in various experiments are reviewed. These models are scored to indicate how comprehensive they are, indicating how well students were able to write them. The results of
that analysis are compared with the results of this one, showing if students who scored well on
the scientific model question on the survey also performed well in writing models during the
course.

It is also possible that this instruction impacted the way students participated in scientific
practice; that is, asking students to learn about and create scientific models during their
chemistry labs may have had an effect on how they designed and executed their own
experiments. This is the focus of Research Question 3 and appears in Chapter 6. Overall, though,
it is clear that gains were made by students in Fall 2017 in understanding the purpose of scientific
models and the molecular nature of chemical models, particularly.
Chapter 5. How the Implementation of a Curriculum with Explicit Instruction on Scientific Models Can Impact Students’ Learning in the Lab

Introduction
The goal of this chapter is to measure students’ learning about the NOS topics of scientific models and model building by analyzing how well students apply this content. In particular, how well can students apply their knowledge of scientific models by writing their own models during an experimental design exercise? The previous chapter showed that, after participating in this curriculum, students in Fall 2017 were better able to define scientific models. The analysis that follows takes this a step further by examining how well students are able to write scientific models. Specifically, how well do these student-constructed models reflect the tenets of what makes a good scientific model? By comparing the performance of students between two different iterations of this curriculum (Fall 2017 versus Spring 2017), this study also attempts to determine the impact of various instructional choices on the effectiveness of the curriculum. In essence, which is more important: the curricular materials or the ways they are approached by different instructors?

Research Question

How well are students able to write scientific models? How important were the instructional choices of the lab instructors and GSIs in impacting students’ abilities to do this?

This question was investigated by looking at data collected from the Chem 1AL (general chemistry laboratory for non-chemistry majors) course in the Spring and Fall semesters of 2017. The data is comprised of students’ written models from the Polymers 1 and Polymers 2 labs (see Appendix 1 for the full text of these labs). Notes from weekly meetings between the course instructor and the graduate student instructors (or GSIs) of that semester are used to provide insight into the instructor’s role in impacting the way GSIs approach NOS material in their own sections. There is also video recordings of the GSIs in an attempt to link students’ performances on those questions with the methods of their GSIs.

Curriculum Materials
The data in this study was collected primarily from the Polymers 1 and Polymers 2 labs (see Chapter 3 for more details on the labs and the curriculum design process). These labs are part of a two-week unit on polymers. The chemistry content that is highlighted includes concepts such as intermolecular forces, crosslinking, molecular shape, structure-property relationships, and property-property relationships. The NOS content (as described in Chapter 2) focuses on scientific models, the process of building and refining models (i.e. generating hypotheses, designing experiments, and interpreting results), and the role of the scientific community in building scientific knowledge.

The Polymers 1 lab is structured to be a guided exploration of polymer properties and the role of models in the scientific process. Part 1 of the experiment gives the students a table of
various combinations of polymeric ingredients to combine and test qualitatively. They track how sticky, stretchy, viscous, and bouncy the resulting solutions are. Often, the GSI will follow this portion of the lab with a discussion of the results and how the molecular-level interactions of the various compounds impact the macroscopic properties of the polymer. The second part of the lab has students quantitatively examining the impact of the ratio of PVA to guar gum on the stretchiness of the resulting polymer. They are also asked to write a model for the stretchiness of a polymer: what causes stretchiness on a molecular level? This is the first time students are prompted to write a model on their own (model P1), although they write several models in the previous experiment as a class when they explore the impact of molecular shape on the sense of smell. This is meant to be a scaffolding exercise to prepare them for next week, when they have much less guidance in designing their own experiments around polymer properties.

The Polymers 2 lab builds on the content knowledge students gained from Polymers 1 by asking them to apply it in a new, less guided context. Students either choose or are assigned a polymer property (i.e. bounciness, stretchiness, stickiness, etc.) to optimize as a toy for children. They must make the bounciest, stretchiest, or stickiest polymer solution they can within certain mass and volume parameters. They are given the same chemicals they used in Part 1 of Polymers 1, and it is their task to design three experiments to help them reach their goals. They are also directed to start with an initial model (P2M1) for their polymer property, in the hopes it will help them design more targeted, thoughtful experiments. They then use their results to refine their model (P2M2) and reflect on the process of model building. Students take their best “polymer recipe” and exchange it with another group to test each other’s results. This helps the students realize how important written communication is in science (i.e. did they write their procedure down explicitly so that someone else could repeat their experiment?). They also come together with other groups working on the same property to discuss measurement methods as a standardization technique. This exposes them to the ideas of debate and agreement in the scientific community, as well. While all these topics are addressed in Polymers 1 and Polymers 2, the following analysis will focus on the learning that occurs around models and model construction by examining the models written by students in Polymers 1 (P1) and the initial and final models they write in Polymers 2 (P2M1 and P2M2). See Appendix 1 for the full text of these labs.

<table>
<thead>
<tr>
<th>Question</th>
<th>Model Writing Context</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>Students write a scientific model for stretchiness. They are given an experiment, with all variables listed, that asks them to determine the impact of the PVA to guar gum ratio on a polymer solution’s stretchiness.</td>
</tr>
<tr>
<td>P2M1</td>
<td>Students write an initial model for the specific polymer property they are optimizing in the Polymers 2 lab. They must design their own experiments around variables of their own choosing. There is much less guidance for writing this model compared to P1.</td>
</tr>
<tr>
<td>P2M2</td>
<td>Students write a revised version of P2M1 after conducting three experiments of their own design.</td>
</tr>
</tbody>
</table>

Table 1. Description of the model writing questions students encounter in Polymers 1 and Polymers 2. They are the focus of the analysis that follows.
Methods

Participants:

Chem 1AL is the laboratory course for general chemistry for nonmajors. The majority of the students in this course identify as life sciences/biology majors, with a significant number of engineering majors and a smattering of other subjects filling out the rest of the population. The exact make-up differs between the fall and spring semesters, particularly with regards to the engineering portion of the population (see Figure 1). The fall semester is also typically larger, with around 1200 students enrolling in the course. The spring semester enrollment is usually around 700 students.

![Figure 1. Distribution of majors in Chem 1AL for spring and fall semesters, 2017](image)

Student Consent:

As prescribed by the IRB for this project, consent was obtained from students via a survey that was distributed electronically through the course website at the beginning and end of the semester. Only students who consented to participate in the study and to allow the research team to utilize their work from the course were included in the analyses presented here. Therefore, the demographics presented here are representative of the consenting students, which amounted to about 550 students from Fall 2017 and about 300 students from Spring 2017. Data was unfortunately not available for the entire class population of either semester.

Participant Background:

Below are figures describing other pertinent demographic information for the subjects in each semester. They indicate that generally speaking, the students in Chem 1AL during the Fall and Spring 2017 semesters were fairly similar with regards to prior chemistry experience, inquiry participation, mother’s education level, and self-reported socioeconomic status.
Figure 2. Previous chemistry experience of study subjects. A score of 0 indicates the student had less than 2 semesters of chemistry prior to Chem 1AL. A score of 1 indicates the student had 2 semesters of chemistry, and a score of 2 indicates the student had more than 2 semesters of chemistry prior to this course.

Figure 3. Inquiry participation by students prior to Chem 1AL. Inquiry here is meant to refer to experiments where students design at least a part of a procedure or experiment. Those students received a score of 1, whereas a score of 0 indicates a student who did not participate in labs or only followed a written procedure given to them by an instructor.

Figure 4. Education level achieved by student’s mother. This is a variable frequently used to represent socioeconomic status. A score of 0 indicates a student whose mother has a high school degree or did not finish high school. A score of 1 indicates a student whose mother completed some college or received a 2-year degree. A score of 2 represents
a student whose mother received a 4-year degree. A score of 3 represents a student whose mother completed some graduate school or received a graduate degree.

![ SES for Fall and Spring, 2017 ]

**Figure 5.** Self-reported socioeconomic status for students in Fall and Spring 2017 Chem 1AL. Note that these numbers represent a slightly smaller population than the earlier figures because students can decline to report this value. A score of 0 represents a student reporting that they grew up in a low-income/poor/working class home. A score of 1 represents a student reporting that they grew up in a middle-class home. A score of 2 indicates a student identifies their childhood household as upper-middle class/professional middle class/wealthy.

**Instructor and Graduate Student Instructors (GSIs):**

The course was taught by one instructor per semester, with Fall and Spring 2017 being taught by different instructors. The instructor conducts a one-hour lab lecture that all students must attend once a week. The students also attend a three-hour lab section once a week that is taught by a GSI. Each section has approximately 25-30 students. Almost all sections also have a Teacher Scholar, an undergraduate student who has taken Chem 1AL before and can act as a peer mentor to the students and an aide to the GSI, who is often new to teaching this course.

Data was also collected on the GSIs (graduate student instructors) who taught the individual lab sections. Each GSI taught two sections. Video was taken of consenting GSIs while teaching, and audio of their interactions with students was captured throughout the three-hour class period. This data was collected for 8 of the 20 GSIs teaching in Fall 2017, representing 16 of the 40 lab sections that semester. Exploratory video data was also collected for the Spring 2017 GSIs to determine how they approached NOS topics in general.

**Data Collection:**

Below is a table detailing the data collected for this study. Surveys were given at the beginning and end of each semester, gathering demographic information, details of students’ past experiences with chemistry, and answers to course evaluation questions. This is also where students were given the opportunity to consent to participating in the study. Student work was collected via the Chem 1AL course website. Here, students uploaded pdf’s of the report sheets that they filled out at the end of each lab. These reports included answers to questions about pertinent chemistry content as well as opportunities to apply and reflect on topics concerning the nature of science. Video was taken of consenting GSIs during several lab sessions throughout the semester in which they were teaching. This provided an opportunity to examine how GSIs had discussed nature of science content in the classroom. Audio data was collected in Fall 2017.
in the hopes of providing information about the frequency and types of interactions GSIs had with students concerning the nature of science. Finally, GSI meetings, which occur weekly with the lab instructor, were attended by myself in Fall 2017 to track how the instructor communicated with the GSIs with regards to the value of the nature of science and potential teaching methods for approaching the subject with their students.

<table>
<thead>
<tr>
<th>Semester</th>
<th>Pre-Survey</th>
<th>Post-Survey</th>
<th>Audio</th>
<th>Video</th>
<th>Student Work</th>
<th>GSI Meeting Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spring 2017</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td>√</td>
<td></td>
</tr>
<tr>
<td>Fall 2017</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td></td>
<td>√</td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Data sources collected in Fall and Spring 2017 Chem 1AL

Data Analysis: Surveys

The pre- and post-course surveys were used in this study to obtain demographic data that could be used to control for semester and section differences in the study population. There was no way to control for the make-up of the course or the sections—students had to be able to choose the lab time that worked best with their schedules. The demographic data used in this study included information on the number of semesters of chemistry the students had taken prior to this course, their prior experience with inquiry activities, and mother’s education level and students’ self-reported household income to account for their socioeconomic status. Prior chemistry experience has been shown in previous assessment work to be predictive for student performance in Chem 1AL, particularly with regards to final grades in the course. The surveys were also used to collect answers to questions about scientific models to help determine if there were changes in students’ understanding of this idea over the course of the semester (see Chapter 4).

Data Analysis: Video

The video was used to study consenting GSIs and their classroom methods during the Polymers 1 and Polymers 2 lab periods. How did they address scientific models in whole class discussions? How much time did they devote to models and model construction during pre-lab lectures and mid-lab discussions? After watching footage for both Polymers 1 and Polymers 2 for eight GSIs, they were ranked according to how much they discussed the writing and development of chemical models. A GSI received a ranking of 0 for no mention of these concepts, a ranking of 1 for mentioning them during either Polymers 1 or Polymers 2, or a ranking of 2 for discussing these topics in both lab periods.

Data Analysis: Student Work

The student work used for this study included the report sheets students turned in for Polymers 1 and Polymers 2 (see Appendix 1). To briefly recount the information found in the Curriculum Materials section, in Polymers 1 students must complete a prescribed experiment to test the impact of the ratio of PVA to guar gum on stretchiness. Every student completes the same experiment, and they are provided the ratios for testing. However, they are also asked to provide a written model that they are testing with that experiment. This is P1. Similarly, students
must write down initial models for a polymer property in Polymers 2 (P2M1), in the hopes that this will encourage them to make more effective experimental design choices during the lab. They then conduct their experiments and write a final model (P2M2) that should reflect their findings from these experiments.

To measure how well the students are able to apply their new knowledge of models, their answers to these three questions were coded to reflect what indicates a “good model.” The coding scheme was developed by considering the curriculum and how it presents the idea of a model along with the answers students were providing to the questions. The final version of this scheme was the result of several iterations, multiple rounds of coding, and many conversations with members of the research team. Students provided a range of answers: from experimental plans and hypotheses on the low end (with no mention of molecular level activity) to logical connections between the molecular and macroscopic level on the high end. The coding scheme used is as follows:

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No model; plan only; no connection between molecular level and macroscopic level</td>
</tr>
<tr>
<td>1</td>
<td>Hypothesis or prediction; connection between polymer recipe and macroscopic level; no connection between molecular level and macroscopic level; potential vague mention of molecular level</td>
</tr>
<tr>
<td>2</td>
<td>Vague connection between molecular and macroscopic levels; no predictive power</td>
</tr>
<tr>
<td>3</td>
<td>Clear connection between molecular and macroscopic levels; vague or inconsistent molecular level explanation; predictive power</td>
</tr>
<tr>
<td>4</td>
<td>Logical or relevant connection between molecular and macroscopic levels; must include reference to both molecular activity and structure/shape; predictive power</td>
</tr>
</tbody>
</table>

Table 3. Coding scheme for model questions in Polymers 1 and Polymers 2. Answers receiving a score of 0, 1, or 2 are not considered models, whereas an answer with a score of 3 or 4 is a chemical model, in that it explains the macroscopic phenomenon on a molecular level and is predictive of specific cases.

The reliability of the coding scheme was determined through the use of two unique raters. After a training session, 50 coded responses from Fall 2017 were compared using squared weights to account for the large breadth of answers (i.e. two raters scoring the same response as a 3 and a 4 is closer than two raters scoring the same response as a 2 and a 4, etc.). The kappa for this analysis was 0.81.

Data Analysis: Course-wide Performance Comparison of Spring 2017 vs. Fall 2017

Two analyses follow from this. First, the Fall 2017 Chem 1AL experience can be compared with that of Spring 2017. The most significant differences between these two semesters were the make-up of the student population and the instructor of the course. Because these population differences can be controlled for with the demographic data collected from the survey, it is possible to determine the role of the instructor in impacting the performance of the students on these questions. This can be determined by analyzing a hierarchical model of the data, as the course structure that clusters several students with one GSI lends itself well to a two-level model (see Figure 6).
The equations for this analysis follow this pattern:

\[
y_{ij} = \beta_0 + \beta_1 \text{chemsem}_{ij} + \beta_2 \text{mothersed}_{ij} + \beta_3 \text{spring}_j + \zeta_j + \epsilon_{ij}
\]

\[
\zeta_j | x_{ij}, w_j \sim N(0, \psi)
\]

\[
\epsilon_{ij} | \zeta_j, x_{ij}, w_j \sim N(0, \theta)
\]

where \(y_{ij}\) is the expected outcome for performance of student \(i\) with GSI \(j\) on any of the three model questions, \(\beta_1\) is the coefficient for the prior chemistry experience of the student, \(\beta_2\) is the coefficient for the level of education of the student’s mother, and \(\beta_3\) is the coefficient for the student’s semester, i.e. whether they were in Spring or Fall 2017. Also, \(x_{ij}\) represents level 1 covariates, \(w_j\) represents level 2 covariates, \(\zeta_j\) is the level 2 residual and \(\epsilon_{ij}\) is the level 1 residual.

With this analysis it was assumed that the distributions of the error terms were normal, mean-centered, and uncorrelated (see Appendix 2 for more details). The coefficients indicate the impact of these variables on a student’s performance on a given problem, and the p-values determine if that variable is statistically significant in determining that performance.

**Data Analysis: Section-specific Analysis of Fall 2017**

Second, the impact of section differences on the performance of students can be investigated by comparing the coding results for different sections from Fall 2017. It is possible to determine if, in fact, there were differences based on section. This was determined through ICC (intraclass correlation) values which were calculated as a measure of the significance of the role of sections and/or GSIs in influencing students’ performances on these model writing questions. The first step for reaching these ICC values involved a variance component model. This includes no covariates, and follows this pattern:

\[
y_{ijk} = \beta_0 + \zeta_k + \zeta_{jk} + \epsilon_{ijk}
\]

\[
\zeta_k | x_{ijk}, w_k \sim N(0, \psi^3)
\]

\[
\zeta_{jk} | \zeta_k, x_{ijk}, w_k \sim N(0, \psi^2)
\]

\[
\epsilon_{ijk} | \zeta_{jk}, \zeta_k, x_{ijk}, w_k \sim N(0, \theta)
\]

where \(y_{ijk}\) is any outcome variable for a given student in section \(j\) with GSI \(k\). The idea behind this analysis is that it can determine if there is a quantitative reason to investigate section or GSI differences as a source of variance in outcome scores. In other words, without thinking about what the cause of the variance between sections or GSIs might be, is it justifiable to account for the three-level structure of the course? The ICC values that were calculated from this analysis
indicate the significance of the role of sections/GSIs on students’ abilities to write scientific models.

Following that, if there is a significant section-based effect, an analysis of a three-level model that accounts for demographic and section-level differences can be conducted. This assumes that students in the same section or who had the same GSI most likely had scores that were more closely correlated to each other than students in different sections or with different GSIs.

The equations for this analysis follow this pattern:

\[
y_{ijk} = \beta_0 + \beta_1 \text{chemsem}_{ijk} + \beta_2 \text{mothersed}_{ijk} + \beta_3 \text{stretchiness}_{ijk} + \beta_4 \text{GSIrank}_k + \zeta^{(2)}_{jk} + \zeta^{(3)}_k + \epsilon_{ijk}
\]

- \( \zeta_k \mid x_{ijk}, w_k \sim N(0, \psi^3) \)
- \( \zeta_{jk} \mid \zeta_k, x_{ijk}, w_k \sim N(0, \psi^2) \)
- \( \epsilon_{ijk} \mid \zeta_{jk}, \zeta_k, x_{ijk}, w_k \sim N(0, \theta) \)

where \( y_{ijk} \) is any of the model scores for a student in section \( j \) with GSI \( k \); chemsem, mothersed, and stretchiness are all control variables for a given student representing his or her prior chemistry experience, mother’s education level, and chosen property for the Polymers 2 experiment; and GSIrank represents the performance of that student’s GSI with regards to models and model construction as determined by the video data of their instruction methods in the laboratory. \( \zeta_k \) is the residual for level 3, \( \zeta_{jk} \) is the residual for level 2, and \( \epsilon_{ijk} \) is the residual for level 1. \( x_{ijk} \) represents level 1 covariates, and \( w_k \) represents level 3 covariates. All random and residual terms were assumed to be distributed normally with no correlation to each other. Note that stretchiness is a dummy variable, where a value of 1 indicates that the student optimized stretchiness in Polymers 2, and a value of 0 means that student optimized a different property. This was important to control for as stretchiness was the property studied in Polymers 1; therefore, students who worked on that property again in Polymers 2 may have gotten a head start on thinking of appropriate models for stretchiness, so to speak, inflating their scores compared to students working on other properties. The coefficients of all these terms indicate the weight of their impact on a student’s performance on these model questions, and the p-values from the analysis indicate the statistical significance of these variables.
Results and Discussion

Spring 2017 P1/P2M2 Results:

<table>
<thead>
<tr>
<th>Outcome Variable</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>1.15</td>
<td>1.08</td>
</tr>
<tr>
<td>P2M2</td>
<td>1.07</td>
<td>1.02</td>
</tr>
<tr>
<td>P2M2</td>
<td>1.22</td>
<td>0.89</td>
</tr>
</tbody>
</table>

Table 4. Descriptive statistics for outcome variables for Spring 2017. n=274 for Spring. Minimum value of 0 for all variables, and maximum value of 4.

<table>
<thead>
<tr>
<th>Semester</th>
<th>Variable</th>
<th>Mean</th>
<th>Standard Error</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spring</td>
<td>P1</td>
<td>1.15</td>
<td>0.07</td>
<td>1.08</td>
</tr>
<tr>
<td>Spring</td>
<td>P2M2</td>
<td>1.22</td>
<td>0.05</td>
<td>0.89</td>
</tr>
</tbody>
</table>

\[ t=1.01, \text{df}=273 \]
\[ p=0.16 \]
\[ \text{Cohen’s } d=0.05 \]

Table 5. Results of a paired, two-tailed t-test comparing the mean model scores on P1 and P2M2 for students in Spring 2017 (n=274). This indicates students in Spring 2017 did not make gains in their ability to write models.

Spring 2017 was the first iteration for this version of the curriculum. A paired, two-tailed t-test was conducted to determine if students made any improvements in their ability to write models between the beginning and end of the Polymers unit. This is not a strict pre/post analysis; while P1 was the first occasion for students to write their own models, it is not strictly the same question as students encounter for P2M1 or P2M2 (see Curriculum Materials section). Therefore, this analysis shows how students progressed in their ability to write models over the course of the Polymers unit. The high p-value for Spring 2017 in Table 7 means there was no significant difference in the performance of Spring students over the course of the Polymers unit. Students did not show any improved ability to write scientific models over the course of the Polymers unit.
Fall 2017 P1/P2M2 Results:

<table>
<thead>
<tr>
<th>Outcome Variable</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>2.16</td>
<td>1.23</td>
</tr>
<tr>
<td>P2M1</td>
<td>2.21</td>
<td>1.29</td>
</tr>
<tr>
<td>P2M2</td>
<td>2.34</td>
<td>1.14</td>
</tr>
</tbody>
</table>

Table 6. Descriptive statistics for outcome variables for Fall 2017. n=536. Minimum value of 0 for all variables, and maximum value of 4.

<table>
<thead>
<tr>
<th>Semester</th>
<th>Variable</th>
<th>Mean</th>
<th>Standard Error</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fall</td>
<td>P1</td>
<td>2.16</td>
<td>0.05</td>
<td>1.23</td>
</tr>
<tr>
<td></td>
<td>P2M2</td>
<td>2.34</td>
<td>0.05</td>
<td>1.14</td>
</tr>
</tbody>
</table>

$t=-2.74$, df=535  
p=0.0032  
Cohen’s $d=0.15$

Table 7. Results of a paired, two-tailed t-test comparing the mean model scores on P1 and P2M2 for students in Fall 2017 (n=536). This indicates students in Fall 2017 did make small but significant gains in their ability to write models over the course of the Polymers lab series.

The p-value of the same test performed for Fall 2017 was less than 0.05, meaning that the difference in means for P1 and P2M2 was significant, though very small (as evidence by the low Cohen’s d value). The narrow difference between P1 and P2M2 is not concerning, however, as the contexts of the model writing exercise for these two time periods were very different. P1 was the first time students wrote their own models, which is a difficult skill to master. The curriculum accounts for this, though, by providing much scaffolding for this activity. Students are not designing their own experiments to test—they are given the experiment in this exercise, so it can act as a sort of hint for what should go into their models. They were exposed to the model building process in the previous lab, so it is also not an entirely new subject for them. In Polymers 2, though, there is far less guidance for the students. They must devise their own hypotheses and experiments from a model of their own conception; none of these components of the scientific process is provided for them to help their model writing process. Writing a model from scratch is a more difficult activity than writing one for a given experimental procedure, so the fact that the Fall 2017 class showed any improvement in their models from P1 to P2M2 is impressive, even though the difference in the means seems small.
Validation of Findings from Chapter 4 for Fall 2017:

<table>
<thead>
<tr>
<th>Mean RQ 1 Model Score</th>
<th>Mean P2M2 Score</th>
<th>Standard Error</th>
<th>N</th>
<th>p-value: Medium</th>
<th>p-value: High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than 2</td>
<td>2.16</td>
<td>0.12</td>
<td>89</td>
<td>0.21</td>
<td>0.0005*</td>
</tr>
<tr>
<td>Equal to 2</td>
<td>2.28</td>
<td>0.09</td>
<td>166</td>
<td>--</td>
<td>0.001*</td>
</tr>
<tr>
<td>Greater than 2</td>
<td>2.64</td>
<td>0.08</td>
<td>178</td>
<td>0.001*</td>
<td>--</td>
</tr>
</tbody>
</table>

Table 8. Mean P2M2 scores for low-, medium-, and high-performing Chem 1AL Fall 2017 students on the post-course survey question, “In your own words, what is a scientific model?” This is referring specifically to the overall sum of a student’s scores in each possible category. This table also reports the p-values for unpaired, two-tailed t-tests between the various performance groups. The values with an asterisk indicate a significant difference between the P2M2 averages of those performance groups.

The findings in the above section are most useful when comparing them to students’ answers to the scientific model question on the post-course survey, as analyzed in the previous chapter. As seen in Table 6, students who scored higher on the RQ 1 scientific model question did write better models at the end of the Polymers unit on average. The high-performing group had a significantly higher mean P2M2 score than either the low- or medium-performing group. Thus, these coding schemes are able to discern higher levels of understanding of scientific models: one by measuring students’ ability to define scientific models and the other by measuring their ability to apply this knowledge by writing their own models.
Comparison of Fall 2017 to Spring 2017:

<table>
<thead>
<tr>
<th>Outcome Variable</th>
<th>Covariate</th>
<th>Coefficient</th>
<th>p-value</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>prev_chem</td>
<td>1</td>
<td>0.21</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>0.17</td>
<td>0.29</td>
</tr>
<tr>
<td>mothers_ed</td>
<td>1</td>
<td>-0.05</td>
<td>0.74</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.09</td>
<td>0.47</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.12</td>
<td>0.32</td>
<td>0.12</td>
</tr>
<tr>
<td>spring</td>
<td></td>
<td>-0.96</td>
<td>&lt;0.0005</td>
<td>0.18</td>
</tr>
<tr>
<td>chi=0.0079</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P2M1</td>
<td>prev_chem</td>
<td>1</td>
<td>-0.03</td>
<td>0.84</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>0.05</td>
<td>0.75</td>
</tr>
<tr>
<td>mothers_ed</td>
<td>1</td>
<td>-0.04</td>
<td>0.77</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.11</td>
<td>0.38</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.18</td>
<td>0.12</td>
<td>0.12</td>
</tr>
<tr>
<td>spring</td>
<td></td>
<td>-1.06</td>
<td>&lt;0.0005</td>
<td>0.18</td>
</tr>
<tr>
<td>chi&lt;0.00005</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P2M2</td>
<td>prev_chem</td>
<td>1</td>
<td>-0.12</td>
<td>0.37</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>-0.03</td>
<td>0.81</td>
</tr>
<tr>
<td>mothers_ed</td>
<td>1</td>
<td>-0.27</td>
<td>0.04</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>-0.12</td>
<td>0.28</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>-0.003</td>
<td>0.98</td>
<td>0.11</td>
</tr>
<tr>
<td>spring</td>
<td></td>
<td>-1.07</td>
<td>&lt;0.0005</td>
<td>0.15</td>
</tr>
</tbody>
</table>

Table 9. Analysis of impact of semester on model question outcomes for Spring and Fall 2017 for P1 (Polymers 1 model question), P2M1 (Polymers 2 initial model question), and P2M2 (Polymers 2 final model question). This was done through a hierarchical model analysis with n=810 students (536 from Fall 2017 and 234 from Spring 2017) nested in 37 groups (i.e. 37 different GSIs), further nested in two semesters (Spring and Fall). There was a minimum of 6 observations (students) per group (GSI) and a maximum of 55. The average number of observations per group was 21.9. Variable names can be found in Data Analysis section.

The hierarchical model analysis shows that even controlling for population differences between semesters, there is a significant difference in performance on all model questions, with students in Fall 2017 consistently predicted to outperform students in Spring 2017 by more than a full point on a four-point scale, on average. The control variables are not significant (p>0.05), and this is true when the analysis is conducted with other demographic variables, as well. Inquiry level is not predicted to significantly impact student performance, and SES has similar results using either students’ self-reported income level or mother’s education level as the representative variable. Therefore, the second iteration of this curriculum and its implementation occurring in Fall 2017 was more successful than the first iteration in Spring 2017.

What caused this difference between iterations, though? Why would the Fall 2017 student population perform significantly and consistently better than that of Spring 2017? Answering this question must begin by examining what was altered between the two semesters.
First, the student populations were not the same. This was accounted for in the regression by including demographic control variables, but it is a small validity concern. Second, there were some minor differences in the materials; the arrangement of the questions in Polymers 1 was changed in Fall 2017 to make the flow of ideas more natural (see Appendix 1). However, the substance of the questions, and of Polymers 2, was not impacted. Third, there were different GSIs teaching each semester (with a very small number of exceptions). However, there were not egregious differences in the GSI pool for each semester (i.e. Spring GSIs did not have a significant lack of experience compared to Fall GSIs), and the presence of different GSIs was accounted for with the two-level model (essentially, controlling for that variable). Finally, the biggest difference between Spring and Fall 2017 was instructors. The laboratory course instructors were different between Spring and Fall, suggesting that the actions of the instructors were a significant determinant of students’ performances on model questions.

A simplified profile of each instructor follows. The instructor for Chem 1AL in the Spring was a first-time instructor finishing his Ph.D. in chemistry. He was very willing to work with the research group in using this curriculum and allowing the members to gather data. However, it was clear from meetings with him, as well as his weekly GSI meetings, that he considered chemistry content and experimental skills to be the most important learning outcomes for the course, rather than explicit NOS topics. As a consequence, the GSIs that semester rarely discussed these topics with their students in the lab setting, and there was no emphasis placed on learning how to write or construct models. The final exam did not include any questions related to these topics. It was clear that the instructor did not value model building as a learning outcome.

The opposite was true for Fall 2017, however. The instructor of the course has been heavily involved in the development and implementation of this curriculum since the beginning. Therefore, she was very receptive to not only using the curriculum but working with me to implement it in such a way as to help students learn about these NOS topics. For example, I was invited to attend her weekly GSI meetings where I could tell the GSIs directly what the goal of the curriculum is with scientific models, particularly during the weeks of the Polymers labs. She, I, and GSIs who had taught the lab before were able to give advice on techniques for broaching these topics in their pre-lab lectures and with students individually or in small groups. The lab instructor even discussed these topics in her weekly lab lecture, which all students must attend. Therefore, she made it clear to her students and her GSIs that she valued this learning outcome. Consequently, the GSIs almost universally mentioned topics of models and model construction in their lab settings.

These observations, coupled with the quantitative analysis above, suggest that the instructor plays a very important role in impacting students’ learning and application of NOS topics such as model and model building. It is not only his or her direct instruction on the topic that matters; it is also how the instructor directs the GSIs (or TAs) to implement the curriculum in the lab itself that determines how well students perform on these types of questions. That direction could be implicit (i.e. how the Spring 2017 instructor never mentioned the importance of models to his GSIs) or explicit (i.e. how the Fall 2017 instructor took time to go over methods for teaching NOS topics in the laboratory). After all, the GSIs all taught from essentially the same materials, with all the same background information given to them and the students about the nature of science and with all the same questions that addressed the topics of models and model
construction. There must be a reason none of the GSIs in Spring 2017 addressed this topic in whole class discussion with their students while nearly all the GSIs in Fall 2017 did.

On a final quantitative note, there is one last measure from the hierarchical models above that bears mentioning. The chi values output by the model are all below 0.05. That validates the choice to treat the data as a two-level model, suggesting that it is significant not only that students are in different semesters but also that they are in different sections with different GSIs. A deeper analysis of the data from Fall 2017 allows for an investigation to determine, first, if students of different GSIs performed significantly differently from each other, and second, how the teaching methods of GSIs might lead to improved model writing by the students.

Section-specific Analysis of Fall 2017:

| Model Question | ICC (Section|GSI) | ICC (GSI) | Chi     |
|----------------|------------------------|------------------------|--------|
| P1             | --                     | --                     | 0.48   |
| P2M1           | 0.19                   | 0.18                   | <0.00005 |
| P2M2           | 0.18                   | 0.15                   | <0.00005 |

Table 10. Intraclass correlations for hierarchical model analysis of Fall 2017 outcome data. n=536 students clustered in 40 sections, taught by 20 GSIs (2 sections per GSI). The minimum observations per GSI is 21, and the minimum per section is 9. The maximum observations per GSI is 31, and the maximum per section is 18. The average observations per GSI is 26.8, and the average per section is 13.4.

The chi values in the table above suggest that for the Polymers 2 model questions, sections and GSIs play a significant role in impacting students’ outcomes. However, the chi value for P1 is greater than 0.05, suggesting that there is no significant impact on scores from sections or GSIs. This is an interesting finding. Why would one model question be more dependent on GSIs than another? P1 is an exercise that comes at the end of an experiment that has been provided for the students. Everyone does the same activity, ends up with similar data, and comes to similar conclusions about stretchiness. While asking students to write a model is not traditional, everything else about that experiment is, and asking students to write this model is akin to asking them what the chemistry behind stretchiness is. They are even provided the variables they will be manipulating. Polymers 2, on the other hand, is a high-inquiry activity. Students are asked to write a model, just like in Polymers 1, but there is much less guidance as to what the model should contain. Students do not all work on the same property, and there is no prescribed experiment. They have to come up not only with the chemistry that produces their property but also the variables that could impact it. The outcome of this analysis suggests the GSI is particularly impactful when students write models in high-level inquiry activities, such as experimental design labs, as opposed to more scripted laboratory activities.

Given this finding, it makes sense to examine the methods of the GSIs in various sections from Fall 2017 during the P2 lab in an attempt to see what makes one GSI more effective than another.
Here, there was an attempt made to quantify the impact of the GSIs by ranking their performances during the Polymers 1 and Polymers 2 labs, as analyzed from the video data collected, with the GSI_rank covariate. A value of 0 correlates with a GSI who never discussed models or model construction with the whole class in either Polymers 1 or Polymers 2. A value of 1 indicates a GSI who mentioned these topics in 1 of the 2 lab sessions, and a 2 recognizes that the GSI mentioned these topics in both Polymers 1 and Polymers 2 discussion with their sections. Demographic data was also included to account for differences in the student populations of each section. The stretchiness variable controls for the property investigated by the students; those who experimented with stretchiness had worked on that property during Polymers 1 and, therefore, could have had an advantage in model writing compared to students working on a new, previously unconsidered property.

The results indicate that none of these covariates are significant. The fact that previous chemistry experience and mother’s education level are not significant is interesting, as these variables are often predictors of grades for students in this course. That suggests that this type of high-level inquiry question decreases those demographic gaps. However, the fact that the GSI_rank variable was also not significant suggests that the role of the GSI is more complicated than simply spending time discussing these topics of models and model construction. How they discuss it may be just as important. It is important to note, however, that several of these sections were represented by only a small sample of the students enrolled. It is possible that an analysis conducted with a larger sample size by section would yield different results. Additionally, the video did not capture individual interactions between the GSI and his or her students. Those

---

Table 11. Results of three-level hierarchical model to determine impact of GSIs from Fall 2017 on students’ model question performance. There are 216 students in 16 sections with 8 GSIs for this data set, with a minimum of 25 students per GSI and 9 students per section, a maximum of 31 students per GSI and 18 students per section, and an average of 27 students per GSI and 13.5 students per section. The smaller sample size is a result of the video data; useable video was collected for 8 of the 20 Fall 2017 GSIs.
interactions can play a very significant part in determining how students construct their models. That data may be extracted from the audio data gathered during the study. Evaluating it to determine how often GSIs interact with students in one-on-one or small group situations, along with how GSIs talk about models in those moments, could be key in determining the appropriate implementation of the curriculum for GSIs in the laboratory.

Conclusions and Future Work

The results of the analyses in this study have important implications for the implementation of a curriculum focused on scientific models and model building. First, the students in Fall 2017 do show small gains in their understanding of models. This was measured by analyzing their ability to apply this knowledge by writing models, and it couples with the findings of the previous chapter, which showed similar small but significant gains in this population when measuring understanding of scientific models with a direct survey question. Combining the findings of these two different measurement methods strengthens the assertion that the students of Fall 2017 have improved in their understanding of scientific models.

Second, given the low mean values for the models written in Spring and, to some extent, Fall 2017, writing chemical models appears to be difficult for general chemistry students. It is a new skill for many students, at least utilizing this language and context. However, there are ways to help students write acceptable models. It is clear from the higher, improving scores of Fall 2017 as compared to the lower, stagnant scores of Spring 2017 that this curriculum requires an instructor who shows attention to this subject. Doing this shows the students that value is placed on this topic, and that impacts their performance. Furthermore, it shows the GSIs that this learning outcome is important, encouraging them to spend time discussing this subject with their students. It is simply not enough to have this topic appear in the written materials of the laboratory, even in the assessment portion, if well-written models are one of the goals of the course. Attention to the subject must be given by the instructor.

The role of the GSI is also crucial, as seen in the results of the section-based analysis, particularly for less guided, more independent lab work. However, the impact of the GSIs is not yet fully understood. Further analysis is needed to determine which specific teaching methods are most effective for aiding students in constructing chemical models. It is possible that individual or small group interactions with the GSI while students are planning their experiments and writing their models are the most important contributions the GSI makes in the classroom.

It is important to note the limitations of this analysis. While data from a sizable sample of the population was able to be analyzed for each semester, these findings are still only representative of approximately 40% of Fall 2017 and 30% of Spring 2017. There was no way to control the make-up of these proportions, either, as students participated in the study on a voluntary basis. The populations of each semester are different, though this was controlled for as much as possible in the quantitative analysis. Finally, it must be noted that both the curricular materials and the instructors were different for each iteration of the course. Thus, the analysis determines that the combination of the materials and instructor of Fall 2017 were the source of the success of that iteration, though the fact that the material changes were slight is suggestive of the importance of the instructor and his or her communication to the GSIs about how to approach the topic of scientific models.
Chapter 6. The Impact of Explicit Instruction on Scientific Models on Students’ Scientific Practices

Introduction

Can a curriculum that highlights a focus on the purpose and execution of scientific research impact the way students perform in the laboratory? This study answers this question by assessing a curriculum that includes explicit instruction on the topics of scientific models and model building in the context of the chemistry laboratory. The goal of the curriculum was two-fold: help students learn about these topics and encourage them to apply these lessons in a setting of scientific practice. It was hoped that explicit instruction on models and model building could act as a scaffold for more difficult experiments that have less guidance for the students. This study specifically examines students’ performances during an experimental design activity in an introductory chemistry laboratory course to determine if this type of instruction could help them to be more thoughtful and successful experimentalists during a challenging, high-level inquiry laboratory. Data from two separate semesters (Fall 2017 and Spring 2017) will be compared to assess two different iterations of this curriculum.

Research Question

How did this curriculum impact students’ experimental design approaches in a high-inquiry laboratory experience? Specifically, did the requirement for students to write an initial and final model before and after conducting their own experiments lead to logical, thoughtful choices in their experimental design?

One of the ideas behind including model writing in the curriculum was to help students as they engage in experimental design. It is a challenging assignment, particularly for students in their first university-level chemistry course. Their chemistry content knowledge, as well as their lab content knowledge, is somewhat limited. However, it was hoped that the appropriate scaffolding would enable the students to engage in this type of high-inquiry activity. Experimental design is an important part of chemistry practice, and one that also applies to other sciences. That makes it desirable to include in a course for non-chemistry majors. The aim of this curriculum adjustment was to prompt students to write an initial model for their property before designing any experiments to help them make thoughtful, informed choices in their designs, rather than simply choosing variables at random to test, as they had been known to do in the past. This study attempts to determine if this model-based scaffolding helped students with the experimental design process.

Curriculum Materials

This study focuses specifically on data collected from the Polymers 1/Polymers 2 series. This is a two-week set of labs that introduces students to polymers and some of the chemical concepts involved in their formation and macroscopic behaviors. In Polymers 1 students qualitatively assess the characteristics of several types of polymer solutions. They also conduct a prescribed experiment to see how the stretchiness of a polymer solution can be impacted by the
change in the ratio of PVA to guar gum. This teaches them about dependent and independent variables. It is also used as a chance to have them write a scientific model to explain what idea they are testing on a molecular level. All of this is scaffolding to prepare them for Polymers 2.

In the second week of this series, students are asked to optimize a polymer property for creating a child’s toy. They either choose or are assigned a property such as bounciness, stretchiness, or stickiness and must use the materials provided (the majority of which they encountered the week before in their qualitative experiment) to make the most extreme type of toy that they can: the bounciest, stretchiest, or stickiest polymer. This is an experimental design activity; they are given very few guidelines other than a total volume parameter. In an effort to help them with this challenge, students are first asked to write an initial model for their property to help them consider the molecular level picture of what is causing and/or impacting this property. From there, they are asked to design three experiments to help them optimize their property and to use their results to revise their initial model and write a final version. The group with the best toy wins bragging rights for the day. For the full experiments, see Appendix 1.

Part 2: Optimization

2. Which property will you and your partner optimize?

3. What is your initial model for your property? Identify three independent variables in your model that you can test to determine their effects on the property of interest. Hint: the data you gathered in the last experiment in Part 1 will help you with this.

Model 1:

Independent variables:
Figure 1. Relevant portions of the report sheet for Polymers 2. Students write an initial model, identify initial variables planned for experiments, write a hypothesis and gather data for three experiments (space for only 1 shown here), contemplate how this might impact their next experiment, and write a final model and reflect on the process of how they arrived at this conclusion.

This is the version used in Fall 2017. In Spring 2017, the question “How does this conclusion inform your choice for the next experiment?” is not present. Everything else, however, is identical.

Methods

Participants and Consent:

Data for this study was collected in Spring and Fall 2017. Chem 1AL in the fall has a typical class size of approximately 1200 students, with a smaller cohort of 700 students in the spring. The students work in labs of 25-30 students. Consent was obtained from students in the
pre-course survey where they were asked to indicate their willingness to participate in the study, in any observations that may occur, and in the collection of their work from the course for future analysis, as required by the project’s IRB. Just under 600 students consented to the study in Fall 2017, and approximately 300 students consented in Spring 2017. The students in this course are majority life science majors with a significant number of engineers included. There are also some social science, public health, and environmental or natural science students. Unfortunately, demographic data was not available for the full course populations in either semester.

![Figure 2. Distribution of majors for Fall and Spring 2017 Chem 1AL students.](image)

**Participant Experience:**

Most students have completed two semesters of chemistry before Chem 1AL, with many completing more than two. Very few have never taken chemistry before this course. Similarly, almost all students have participated in a lab before, though the majority have never had experience with experimental design. This is all fairly typical of an introductory chemistry course in a university.
Figure 3. Previous chemistry and laboratory experience for Fall and Spring 2017 Chem 1AL students. The left side of the graph indicates the proportions of students who have completed less than 2 semesters of chemistry prior to Chem 1AL (score=0), exactly 2 semesters prior to this course (score=1), and more than 2 semesters prior to the course (score=2). The right side of the graph indicates students’ experiences with labs prior to Chem 1AL. Spring has a little less experience with labs.

Instructor and Graduate Student Instructors (GSIs):

The course was taught by one instructor per semester, with Fall and Spring 2017 being taught by different instructors. The instructor conducts a one-hour lab lecture that all students must attend once a week. The students also attend a three-hour lab section once a week that is taught by a GSI. Each section has approximately 25-30 students. Almost all sections also have a Teacher Scholar, an undergraduate student who has taken Chem 1AL before and can act as a peer mentor to the students and an aide to the GSI, who is often new to teaching this course. The GSIs were different for the two semesters, with only one graduate student teaching during both iterations of the course.

Data was also collected on the GSIs (graduate student instructors) who taught the individual lab sections. Each GSI taught two sections. Video was taken of consenting GSIs while teaching, and audio of their interactions with students was captured throughout the three-hour class period. This data was collected for 8 of the 20 GSIs teaching in Fall 2017, representing 16 of the 40 lab sections that semester. Exploratory video data was also collected for the Spring 2017 GSIs to determine how they approached NOS topics in general.

Student Selection:

The work of 66 students from Fall 2017 and 58 students from Spring 2017 was analyzed for this study. A smaller proportion of the population was required due to the extensive nature of this analysis. These students came from two of the high-performing sections as regards model writing (i.e. these sections consistently scored highly on model coding questions, as determined in Chapter 5) and two low-performing sections for Spring 2017. Fall 2017 reviewed data from two high-performing and three low-performing sections. Three sections were chosen, rather than two, simply to ensure a balanced number of students from each type of section (see Table 1 below). This ensured there would be a range of student work from sections with different GSIs.
As the work in the previous chapter shows, a student’s section (and consequently, GSI) can be a significant factor in his or her performance.

<table>
<thead>
<tr>
<th>Section</th>
<th>Consentning Students: Fall</th>
<th>Consentning Students: Spring</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-performing 1</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>High-performing 2</td>
<td>16</td>
<td>14</td>
</tr>
<tr>
<td>Low-performing 1</td>
<td>9</td>
<td>17</td>
</tr>
<tr>
<td>Low-performing 2</td>
<td>13</td>
<td>11</td>
</tr>
<tr>
<td>Low-performing 3</td>
<td>12</td>
<td>--</td>
</tr>
<tr>
<td>Total</td>
<td>66</td>
<td>58</td>
</tr>
</tbody>
</table>

Table 1. Distribution of consenting students by section.

Data Collection:
Surveys were given at the beginning and end of each semester, gathering demographic information and details of students’ past experiences with chemistry. This is also where students were given the opportunity to consent to participating in the study. Student work was collected via the Chem 1AL course website. Here, students uploaded pdf’s of the report sheets that they filled out at the end of each lab. These reports included answers to questions about pertinent chemistry content as well as opportunities to apply and reflect on topics concerning the nature of science. Finally, GSI meetings, which occur weekly with the lab instructor, were attended by myself in Fall 2017 to track how the instructor communicated with the GSIs with regards to the value of the nature of science and potential teaching methods for approaching the subject with their students.

Data Analysis:
To analyze students’ experimental design techniques, their Polymers 2 report sheets were used as the main source of data (see Appendix 1). Previously, their initial and final models were analyzed (see Chapter 5) to determine how well students were able to construct scientific models. Following that, I wanted to determine if they used their initial models in framing their experimental choices. Did their abilities to write models influence how they designed their experiments? Each student’s entire model building process, from writing an initial model to recording experiments used to test that model to writing a refined final model, was examined to determine answers to the questions in Table 2. Answering these questions provided a good indicator of how students approached the experimental design task.
<table>
<thead>
<tr>
<th>Question</th>
<th>Potential Answers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Did students’ experiments test their initial models?</td>
<td>Yes or No</td>
</tr>
<tr>
<td>Did students use the results of their experiments to frame future</td>
<td>Yes or No</td>
</tr>
<tr>
<td>experimental choices?</td>
<td></td>
</tr>
<tr>
<td>How did students use the results of their experiments to frame</td>
<td>Choices, From Model, Recipe, Predicting Future Results</td>
</tr>
<tr>
<td>future experimental choices?</td>
<td></td>
</tr>
<tr>
<td>Did students’ final models differ from their initial models?</td>
<td>Yes or No</td>
</tr>
<tr>
<td>How did students’ experiments impact their final models?</td>
<td>Recipe or Molecular</td>
</tr>
</tbody>
</table>

Table 2. Questions for determining students’ experimental design processes for the Polymers 2 lab.

For the question, “How did students use the results of their experiments to frame future experimental choices,” the “recipe” category is fairly self-explanatory: students used the results of their experiments to influence the polymer recipes they used in future experiments. For example, there were several instances of a student using the first experiment to determine an ideal ratio of certain polymer ingredients to then use in the next two experiments to test for mechanical factors, such as ideal stirring rate or resting time. Sometimes, a student determined a chemical was not good to use in the recipe and moved onto other options in the next experiment. These types of connections were not molecular in nature, nor did they connect to any underlying model explaining the property to be optimized.

The “choices” category indicates that a student used the results of an early experiment to inform the choices they made in later experiments. These choices clearly had some underlying logic beyond using recipe results for the next test. For example, a student may have tested a certain polymer ratio in Experiment 1, recognized that one of the polymers was not a good option, and chosen to replace it with a different polymer based on its structural features. This is a recipe choice, to an extent, but it was influenced by a molecular interpretation of the experimental results, perhaps related to the initial model. Here, students based their experimental design choices on interpretations of data from earlier experiments.

The “from model” category indicated students whose experiments were all driven by their initial model but were not obviously connected to each other in any other way. The subsequent experiments did not rely on the results of earlier experiments, but they were all connected through the underlying logic of the model. Finally, the “predicting future results” category refers to the student who used experimental results to predict the results of future experiments, rather than to help them design those experiments. These are both acceptable ways to design experiments, but they did not seem to fit in the other categories.

For the question, “How did students’ experiments impact their final model,” the answer categories were straight forward. For students who did have model changes from Model 1 to Model 2, they could be either positive or negative changes. A positive molecular change means a student wrote a final model that was more specific on a molecular level. Similarly, a positive recipe change means the student was able to report a more specific recipe for their property optimization than in Model 1. A negative change in the molecular category mean the final model was less explicit on the molecular level. The student might even have written an initial model with molecular-level ideas that disappeared by the final model. A negative recipe change meant the student went into less detail about their final polymer recipe in Model 2.
### Question  Interrater Reliability Score

<table>
<thead>
<tr>
<th>Question</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Did students’ experiments test their initial models?</td>
<td>0.93</td>
</tr>
<tr>
<td>Did students use the results of their experiments to frame future experimental choices?</td>
<td>0.00*</td>
</tr>
<tr>
<td>How did students use the results of their experiments to frame future experimental choices?</td>
<td>0.95</td>
</tr>
<tr>
<td>Did students’ final models differ from their initial models?</td>
<td>0.78</td>
</tr>
<tr>
<td>How did students’ experiments impact their final models?</td>
<td>0.93</td>
</tr>
</tbody>
</table>

Table 3. Results of an interrater reliability test (reporting kappa) between two unique raters after two rounds of coding. About half (34) of the Fall 2017 responses were used for this test. Note the high value of all kappa values, with the exception of the second entry. Here, the kappa value is an artifact of the way it is calculated: it is 0 because the actual agreement and expected agreement between raters were equal. However, the expected agreement for this question was over 93%, meaning there was actually very high agreement between the raters for this question. This all provides evidence of the reliability of the coding scheme.

### Results and Discussion

*Did students’ experiments test their initial models?*

The majority of students in Fall 2017 (74%) did have at least one experiment that tested their initial models. This was true for fewer students in Spring 2017 (62%). In many ways the ability for students to do this depended on the scope of their initial models. Students who wrote vague initial models had difficulty creating three experiments to test them. Sometimes, one experiment was all that could be envisioned for the model, and the subsequent experiments were largely unrelated to the model. Conversely, for some students the initial model was so specific that only one experiment could be designed from it. For example, if the initial model was a particular polymer “recipe,” such as specified amounts of PVA and guar gum, there was only one experiment that could be done: create that recipe and test it for the property of interest. The following experiments may perhaps have built on those results, but they were not directly testing the initial model.

Among the students who did not test their initial models with their designed experiments were those whose initial models were not, in fact, models. They were plans for their experiments or methodologies for testing their polymer solutions. If there was no model to test, then it follows that the experiments could not be related to that model directly. Occasionally, students wrote models and essentially ignored them as they proceeded with their experiments. There was no clear indication in these cases that students were connecting their experiments to their models.

These results suggest that the ability of students to write an initial model that is sufficiently complex is crucial in allowing them to create model-driven experiments. A model that is too vague or too specific makes it difficult for students to use them as tools in forming their experiments. However, a majority of students were able to recognize the importance of designing experiments to test their models, and this was true for more students in Fall 2017 than in Spring 2017.
Did students use the results of their experiments to frame future experimental choices? If so, how?

As can be seen in Figure 4, students overwhelmingly used the results of their early experiments to impact the design of their later experiments in Fall 2017. Only four students did not seem to have a connection among their experiments. The connections were categorized largely as “choices” or “recipe.” The other categories included just a few students. The “from model” category for Fall 2017 is for students whose experiments were all driven by their initial model but were not obviously connected to each other in any other way. However, they were certainly connected through the use of a logical model. These are positive outcomes, but it would be preferable for there to be a higher number of students in the “choices” category rather than the “recipe” category, as that would indicate a deeper connection between the experiments.

In stark contrast the students of Spring 2017 overwhelmingly did not use the results of their early experiments to impact their future experimental design decisions. In fact, 60% of them did not. The ones who did fell mostly in the “from model” category. This category still represents students who wrote experiments that connected to the initial model but not to each other. However, for 6 of these students, the connection to the initial model was superficial. These students wrote a model that scored as a 1, meaning it had no molecular level elements, and they listed three variables that they were planning to test. Even in the model there was no underlying connection between experiments, meaning that generally speaking, the students in Spring 2017 designed experiments with no obvious connections to each other or to any underlying chemical principles.
How did students’ experiments impact their models?

Figure 5. Ways students’ models shifted after completing their experiments for Fall and Spring 2017. Blue bars represent negative shifts, meaning Model 2 in these cases was less developed than Model 1. Orange bars represent positive shifts, meaning Model 2 showed some type of advancement from Model 1.

Overall, sixty-two students from Fall 2017 changed their initial models to something different for Model 2. Only four students did not show significant changes. Of those four, one person started with an initial model that was, in actuality, a plan for the experiment and ended with something similar. Thus, this student went from something that was not a model to something else that was not a model. The other three students saw their initial models as confirmed by their experiments and so had no need to write a new model for Model 2.

Of the students who showed changes in their models, the majority showed positive changes. This means that their final models were more advanced than their initial models. Some of these changes were recipe changes, meaning their final model either added in or improved upon an initial idea for what chemicals to combine and how. These changes were not molecular in nature. However, a large number of students did show molecular level improvements in their models, meaning their experiments allowed them to deepen their understanding of the chemistry involved in polymeric structure-property relationships.

A few students showed negative changes in their models, meaning their final models were less advanced than their initial models. A few of these were students who went from an initial model with molecular level ideas to a final model with fewer or none of these ideas, perhaps giving only a final recipe. Two students had a negative recipe change, as well, where the final recipe was less detailed than the original.

Similarly, 52 of 58 students in Spring 2017 showed changes from their initial to final models, and most of these changes were positive. However, the majority of these changes were recipe-based in nature. Very few students made any significant molecular-level changes in their models after conducting their experiments, meaning few students were able to use these experiments to deepen their knowledge of the chemistry at play. Additionally, the results found in Chapter 5 indicated that the Spring 2017 students generally wrote models that did not include any molecular-level ideas. Combining that with the finding here that few students made molecular-level changes in their models suggests that the Spring 2017 cohort was not able to embrace the role of model writing and model building in scientific practice.
Discussion:

Overall, the results for the Fall 2017 paint a positive picture about the effects of this iteration of the curriculum on students’ experimental design abilities. The vast majority of students did use their experiments to test their initial models, showed logical connections between experiments, and showed improvements between Model 1 and Model 2. However, there is still room for improvement. Few students were able to construct three coherently connected experiments that all tested their initial models. Fewer still made molecular level interpretations of their results as they used them to design their experiments and conclude with an improved model that showed deeper understanding of the chemistry behind their polymer properties. Admittedly, these are difficult goals to achieve. However, the data in this study suggests several ways to help make this a reality.

There is evidence here that model writing and model building are separate learning goals for students. There were several students analyzed in this study that could write decent initial models (perhaps a 3 or even 4 on the scale explained in Chapter 5) but did not engage with them during the experiment. Some even wrote worse models after completing their experiments, suggesting they had no true understanding of the purpose of scientific models. Therefore, it is certainly possible for students to write acceptable models without knowing how to use them effectively.

Similarly, some students wrote overly simple models that had little to no molecular level understanding of their properties but were able to execute a fairly coherent experimental design plan. These students used the results of each experiment to inform the next, even making logical choices in their designs. However, they did not know how to relate that to a model, either at the beginning or the end of the lab. Combining this evidence with that from students who could write models well while designing experiments poorly, it becomes clear that these are separate learning goals that need to be addressed as such.

This idea is supported by similar findings in the literature. As Abd-El-Khalick and Lederman explain (2000), early in the work on NOS, researchers realized that NOS topics had to be treated as cognitive learning goals, rather than affective goals, for students to show any gains in understanding (Durkee, 1974). This seems to be true for models and model building, as well. It is not enough to be able to write acceptable models. This does not imply students are capable of using these models correctly while engaging in scientific practice. This is a separate topic that must be addressed explicitly in course materials and instruction.

One way this can be addressed is improving the wording of the prompt in Polymers 2 asking students to consider how their conclusions from an experiment will influence the next one. This could be changed to more closely mirror some of the exercises in the How the Nose Knows experiment (see Chapter 3) which students complete earlier in the semester. In that lab students develop a model for how the sense of smell works on a molecular level. However, they do it as a class in conjunction with their GSIs. They are prompted to revise their initial model after each experiment they do to come to a final model that has been tested several times. Building this same principle into the Polymers 2 lab by changing that prompt to ask students to consider how their results have impacted their understanding of the molecular-level nature of their properties would be a nice recall to the guided model building they have done in the past. It would also help them recognize the roles of experiments and experimental design in model building.
However, this must also be supported by the actions of the GSIs in the laboratory room itself. GSI meetings can be used to help the GSIs find ways to encourage students to engage with molecular level ideas throughout the experiment, not simply as they write their initial and final models. In Fall 2017 the GSIs were warned that they would see random experimental design choices being made by students if they did not encourage them to be thoughtful about how the experiments relate to each other. This warning was clearly effective; many students in this study connected their experiments together, often by using the conclusions of one experiment to create the polymer to use in the next experiment. Video data shows that the GSIs talked about this very idea in the lab. However, it was often only addressed in a very brief sentence about using the results of one experiment for the next. If the GSIs are instructed to spend more time on this, prompting students to engage with the molecular level as they go, the result may be more of these successful experimental design activities where students are able to write more nuanced models at the end of the lab due to the findings of their own experiments. In essence, if the students can be helped to see overall connections between scientific models, model building, experimental design, and the evolving nature of scientific knowledge (i.e. all topics from NOS) through more explicit instruction from the GSIs, perhaps they will better understand the purpose of the model writing exercise, which is to help them make logical choices to refine their models through experimental design.

All of these findings are supported by the results of the analysis for Spring 2017. In contrast to the fall cohort, Spring 2017 students did not exhibit the same promising experimental design techniques. Fewer students tested their initial models with an experiment. Most students did not write experiments that related to each other or to a unifying chemical principle. Most of the experiments that did connect to each other did so only superficially. Why did this group of students behave so differently from Fall 2017? There were two main differences in the curriculum and its implementation from spring to fall. First, spring did not have the question, “How does this conclusion inform your choice for the next experiment?” in the lab manual. Second, the GSIs were not given any guidance in how to aid their students in making thoughtful experimental design choices. Therefore, there was nothing to prompt students, either in the materials or in their implementation, to consider the importance of each experiment to the goal of improving the initial model. Students were not explicitly told to consider their experimental results during the design process, so most of them did not. Therefore, it is clear that a curriculum that explicitly addresses scientific models and their role in knowledge building and experimental design could help encourage students to employ these ideas in their own experimental design activities. However, this is not a sufficient condition; instructional support is also needed to reinforce the importance of these design choices to the students.

Conclusions and Future Work

Overall, this study suggests that the goals of the curriculum with respect to experimental design are being achieved when coupled with sufficient instruction and prompting questions, and these results can be improved with an iterative adjustment of both materials and instruction methods. It is important to note, however, the limitations of this study. Only a portion of the student work submitted each semester was analyzed, due to the intensive nature of the coding scheme. Therefore, the results cannot be generalized to the entire population. The GSIs were
also different between semesters, though there was no evidence to suggest significant differences in the abilities of the two groups.

An expanded analysis that looks at a larger sample of work from Fall 2017 would be a good next step to solidify these findings. In future, it could be valuable to analyze this topic of experimental design from another source of data, such as interviews with the students or even journal prompts, asking them to write about their experiences designing these experiments. Video of a few groups of students engaged in this lab could also be useful in expanding the research team’s understanding of their experimental design process.
Chapter 7. Conclusions and Future Work

Conclusions

The results of these studies indicate that overall, the redesigned Chem 1AL curriculum led to positive outcomes for students when it was coupled with effective implementation strategies. Students in Fall 2017 made positive learning gains in their understanding of scientific models. This was seen in their answers to the survey question, “In your own words, what is a scientific model?” on the pre- and post-course surveys. It was also apparent in their performance on the final model writing question of the Polymers 2 experiment. These two analyses show students made small but significant gains in both their abstract understanding of models and their ability to apply them to scientific contexts.

Students’ abilities to write chemical models was dependent on the instruction they received in the course. In semesters where the instructor was engaged with explicit instruction on models and model building and discussed it both in the lab lecture and in GSI meetings, students performed significantly better than in semesters where the instructor focused on other areas of chemistry. During the semester with the engaged instructor, the specific section in which a student was enrolled had a significant impact on performance. After controlling for various demographic information, the major cause of this difference in performance points to the GSIs. However, it is difficult to determine what aspects of a GSI’s instruction may be the most impactful on students’ abilities to write scientific models.

Students engaging in this curriculum did employ desirable scientific practices when participating in experimental design with an engaged instructor. Specifically, the vast majority of students designed experiments that tested their initial models; they used the results of early experiments to influence the designs of their future experiments; and they refined their initial models to create more detailed final models, which is the goal of practicing scientists and a major achievement for the curriculum. This only happened, though, when the instructor and GSIs actively engaged with students on the topics of models and model building.

Students responded positively to this curriculum. They enjoyed the labs, found the materials clear and useful, and felt the course improved their ability to engage in experimental design. A significant percentage of students discussed NOS ideas as their most valued learning outcomes when asked in an open-ended question on the post-course survey.

These positive outcomes suggest that a laboratory curriculum with explicit NOS instruction can be effective. It can lead to gains in understanding and application of NOS topics. It can also lead to deeper engagement with chemistry topics, as seen in the Polymers 2 lab for the Fall 2017 students. However, this is incumbent upon the implementation of the curriculum. Instructors must engage with these topics explicitly through discussion in order to show their students they value this type of learning and to prompt students to pursue a higher level of understanding. As can be seen with the analysis of student work in Spring 2017, explicit materials about models and model building, even though they are contextualized and prompt student reflection, is not enough to lead to effective application of such topics. It is also worth noting that these positive outcomes all came from a curriculum that was already established and was adjusted to make its implicit NOS instruction explicit. With the appropriate implementation of such a curriculum, it is possible to achieve positive outcomes for students.
Future Work

There are some simple changes to be made to the curriculum that could improve upon the above findings. For example, changing the prompt in Polymers 2 after each experiment that asks students to consider how their results will impact their next experiment could lead to more engagement with molecular-level ideas. It could be changed to more closely resemble the How the Nose Knows activity and prompt students to consider how their results change their initial models. Perhaps this would work best as an additional question placed just before the original prompt. Additionally, helping the GSIs determine more effective ways to approach instruction on experimental design could have similarly positive impacts. Instructing them to go into more detail about how students should be using their initial models as tools for experimental design, and how models and model building are inextricable parts of scientific practice, could lead to better practices by students.

Further analysis on these topics would be an interesting future project. Conducting interviews of students about their lab practices would be illuminating. Perhaps video of them engaging in experimental design would lend further depth to these findings. However, there is also existing, unexamined data from this project that would be useful to analyze in the spirit of these studies. For example, there is audio of GSIs talking with students during the Polymers 2 lab. The secret to the GSIs’ impact on students as they participate in experimental design could be unlocked by listening to it. There are also many survey questions related to NOS topics that could prove useful. In addition, the curriculum addresses many other NOS ideas aside from models and model building. Analyzing its impact on students’ understanding of these other elements of NOS would be a logical next step in this analysis. As this curriculum has been developed in iterations over many cycles of implementation, analysis, and refinement, there is much that can still be done to continue this process and further improve the laboratory curriculum to positively impact students’ understanding of the nature of science and their engagement in the general chemistry laboratory.
References


Matthews, M.R. (2012). Changing the focus: From the nature of science (NOS) to features of science (FOS). In M.S. Khine (Ed.), *Advances in nature of science research: Concepts and methodologies* (pp. 3-26). Dordrecht: Springer.


Appendix 1: Complete Fall 2017 Versions of the Polymers 1 and Polymers 2 Labs
Polymer Properties and Applications
Part 1: Cross-Linking Reactions in Toy Design

Introduction

Research Question
How do the observable properties of polymers change when their molecular structures change?

Learning Goals
The learning goals for this two-week experiment include understanding the structure and behavior of polymers in crosslinking and how the chemical structure affects the macroscopic (observable) properties of the materials. This lab allows you practice designing your own experimental procedure in the context of product design and product development.

In order to fully understand these changes, you observe and what might be causing them, you will develop a procedure for quantitatively observing the changes in one of these properties as it relates to changes in polymer composition. Experiment design and systematic manipulation of variables are important skills for scientists to develop. You will see these ideas again in the next part of the module.

Green Chemistry: The polymers used in these experiments are non-toxic and water-soluble, making them Inherently Safer and allowing us to conduct the crosslinking reactions using Safer Solvents. Additionally, some of the polymers, such as guar gum and cornstarch, come from Renewable Feedstocks. The crosslinking reactions have high Atom Economy with few unwanted by-products.

Science in Practice: In the last lab “How the Nose Knows,” we began discussing scientific models and their construction. This week’s lab allows us to push deeper into these same themes. Models are an integral part of scientific inquiry; in fact, model construction is a main goal of scientific practice. Models are meant to explain the world around us, and science works continually to improve these explanations to make them as accurate and useful as possible. Scientists refine models through various types of testing. You are probably most familiar with the method of using experiments to test hypotheses. A hypothesis is a prediction about your phenomenon of interest based on the current model. It is not explanatory (that’s the model’s job), and it must be testable. In the last lab, you tested Model 1 (a compound’s smell is dictated by its molecular formula) by smelling two different compounds that had the same molecular formula. Your hypothesis for this test was that both molecules should smell the same; however, your experiment yielded data that contradicted this hypothesis. Therefore, you had to refine your model based on the evidence you collected. This led to Model 2 (a compound’s smell is dictated by its 2D chemical structure, particularly the presence of functional groups).
In “How the Nose Knows,” the experiments you ran to test your hypotheses and improve your model were fairly simple: smell strategically chosen compounds to test the predictions your model suggested. What happens if you have to design your own experiment, though? This week’s lab will introduce you to the major components of an experiment in an effort to prepare you for next week’s lab when you must design your own experiment to create a polymer-based toy. Some of the most important things to consider in experimental design are the variables in your model. Sometimes, they are obvious; other times, they seem to be hidden, perhaps because you have not yet accounted for them in your model. The more variables you can identify in your model, the more you can account for them in your experimental design.

Let’s say you want to make the best chocolate chip cookie. You can control the ingredients you put in the bowl to make the batter and their amounts. You can control the temperature of the oven (though the certainty of that temperature is suspect depending on your oven). You may not realize, however, that humidity can play an important role in baking. If you know that humidity is a variable, you can control for that by measuring the humidity in your kitchen and perhaps using a humidifier or dehumidifier to adjust to your preferred level. You could also bake your cookies in multiple kitchens to judge the impact of varying humidity on your product. Knowing your variables allowed you to design a method for your experiment that was more robust in producing reliable data that will refine your model of how to bake the best chocolate chip cookie.

Not only does identifying variables help you design better experiments; neglecting variables that may not be apparent in your model can also lead to incomplete or uninterpretable results. If you bake your cookies in your dorm and then use the same recipe in your lab partner’s dorm, you may get different results. But if you don’t know that humidity is a factor or that ovens can have temperature variability (even though they both say 350°F), you may be at a loss as to how to interpret these results correctly. You may change your potentially award-winning recipe without having the full set of facts at your disposal, leading to an incomplete or incorrect model.

This lab, you’ll be doing experiments to determine relationships among variables that can affect microscopic (crosslinking) and macroscopic (i.e. stretch length, bounciness, etc.) properties of polymers. These relationships will be important for next week, when you must make a model for one of these properties to make the best polymer-based toy. Part 2 of this week’s lab sets up various experiments you can do to test one of these properties, stretchiness. It’s a practical example of how to determine the most important variables in your model, make a hypothesis based on this model, and develop a method for testing it. It also forces you to consider how your methods may directly impact the data you collect for testing and revising your model. Next lab, we’ll see that models can help you determine the most effective experiments to conduct in order to improve your model.

After doing this lab, you should have a better understanding of these ideas:

**Scientific Inquiry:** How can a thoughtful consideration of all variables involved in your experiment lead to more reliable results?
Polymers

Polymers: The Basics

Polymers are molecules with a very high molecular weight, which are made by connecting many small molecules (the “monomer” building blocks) into long chains. Polymers make up a large part of consumer goods, such as plastic bottles and synthetic fibers. There are a wide variety of types of polymers, both naturally occurring and synthetic. Some examples of natural polymers include peptides and proteins, which are made of amino acid monomers; polysaccharides such as cellulose and starch, which are made of sugar monomers; DNA, which is made of nucleic acid monomers; and rubber, which is made of isoprene monomers.

Some examples of synthetic polymers include plastics such as polyethylene, polystyrene, polyamides (such as nylon), and polyesters (such as polyethylene terephthalate). The names of these polymers are derived either from the names of the monomers used to make them (for example, polystyrene is made from styrene monomers), or from the functional groups that are present in the monomer (for example, polyesters and polyamides contain repeating ester and amide functional groups, respectively).

Polymers are classified in a variety of ways. Often, polymers are grouped according to the method of synthesis - addition reaction versus condensation reactions. Addition polymers are formed by a reaction in which monomer units simply add to one another to form a long-chain polymer. Condensation polymers are formed by the reaction of bi- or poly-functional molecules, with the elimination of some small molecule (such as water, ammonia, or hydrogen chloride) as a by-product. Familiar examples of condensation polymers are nylon, Dacron, and polyurethane. Some simple polymerization reactions are represented symbolically below. Polymers can be made from the same monomer units or from different ones to give the desired property.

Monomer → Polymer

Monomers → Copolymer

A polymer with crosslinking
Polymers are also categorized based on their **physical properties**. Materials that can be softened (melted) by heat and re-formed (molded) into another shape are known as thermoplastics. Weaker, non-covalent bonds and intermolecular attractions are broken during the heating. Both addition and condensation polymers can be so classified. Familiar examples include polyethylene and nylon. Materials that melt initially but on further heating become permanently hardened are known as thermoset plastics. They cannot be softened and remolded without destruction of the polymer because the covalent bonds are broken. Chemically, thermoset plastics are cross-linked polymers.

**Crosslinking Polyalcohols with Borax**

In this two-day experiment, you will explore the crosslinking reactions of several polymers in aqueous solution. You will observe the effects that this crosslinking reaction has on the macroscopic (observable) properties of the solutions. During the second week of the experiment, you and your classmates will select certain properties to optimize to create a new crosslinked polymer formulation that could be used in a toy.

The polymers in this experiment are water soluble due to the presence of alcohol (hydroxyl) functional groups along the backbone, which can form hydrogen bonding interactions with water. When borate ions are added to the polymers solutions, the alcohol functional groups interact with borate ions to **cross-link** (see above) the polymer chains together, either by hydrogen bonding or by covalent bond formation or by a mixture of these two modes.\(^1\) The crosslinked polymer solutions have macroscopic properties that are drastically different from those of the simple aqueous polymer solution.

**Reactions of Borate Ions**

1. **Borax** (sodium tetraborate) dissolves in water to provide borate ions:

   \[
   \text{Na}_2\text{[B}_4\text{O}_5\text{(OH)}_4]\cdot 8\text{H}_2\text{O} \rightarrow 2\text{Na}^+ (\text{aq}) + \text{B}_4\text{O}_5\text{(OH)}_4^{2-} (\text{aq})
   \]

   \[
   \text{B}_4\text{O}_5\text{(OH)}_4^{2-} (\text{aq}) + 7\text{H}_2\text{O}(\text{l}) \rightarrow 4\text{[B(OH)}_4]^- (\text{aq}) + 2\text{H}^+ (\text{aq})
   \]

2. The borate ion has **polar-covalent bonds**. This is an interesting case where the formal charge on an atom does not match the partial charge based on bond polarity. Bond dipoles are expressed such that the arrow points from the less electronegative atom to the more electronegative atom. These bond dipoles work together to form partial charges on atoms in the molecule.

\(^1\) *J.* Phys. Chem. B, 2010, 114 (41), pp 13059–13068
3. Borate ions can form **non-covalent hydrogen bonding** interactions with alcohols. Each alcohol or borate ion can act as either a hydrogen bond donor or a hydrogen bond acceptor. A hydrogen bond acceptor is a neutral or negatively charged atom with a free lone pair of electrons. A hydrogen bond donor is a hydrogen atom bound to an electronegative atom such as oxygen, sulfur, nitrogen, or any of the halogens.

![Diagram of non-covalent hydrogen bonding](image)

4. Borate ions can form **covalent bonds** with alcohols, by breaking and forming boron-oxygen bonds.

![Diagram of covalent bonding](image)

**Calculating Mass Ratios in Mixtures**

The solutions used in the lab this week are described as a mass percent. A solution that has a mass percentage of 1% (w/v) was prepared with 1 g of the solute in 100 mL of the solution. For our purposes, the solute is the polymer or the cross-linking agent and the solvent is water. The following example will show you how to calculate the mass ratio of a mixture of polymers.

*You mix 3 mL of a 5% aqueous solution of guar gum and 2 mL of a 2.5% aqueous solution of corn starch in a beaker. What is the mass ratio of guar gum to corn starch in your new mixture?*
First, calculate how many grams of each compound are in the mixture:

\[
3 \text{ mL (solution guar gum)} \times \frac{5 \text{ g guar gum}}{100 \text{ mL solution guar gum}} = 0.15 \text{ g guar gum}
\]
\[
2 \text{ mL (solution corn starch)} \times \frac{2.5 \text{ g corn starch}}{100 \text{ mL solution corn starch}} = 0.05 \text{ g corn starch}
\]

Then, determine the mass ratio of guar gum to corn starch:

\[
\frac{\text{guar gum mass}}{\text{guar gum mass} \, + \, \text{corn starch mass}} \times 100 = \frac{0.15 \text{ g}}{0.15 \text{ g} + 0.05 \text{ g}} \times 100 = 75\% \text{ guar gum}
\]

Thus, the mass ratio of guar gum to corn starch 75:25.
Prelab Questions

1. Define the following terms: **monomer, polymer, copolymer, cross-linking**  Cite your sources of information.

2. Look up the following polymers: **polyvinyl alcohol, polyvinyl acetate, guar gum, and amylopectin**. Find information about the source of each polymer (for example, is it man-made or naturally occurring? What is it made from or where is it found in nature?). Additionally, find at least two different uses for each polymer. Record this information along with citations for the sources of your information in your notes in a table as shown below.

<table>
<thead>
<tr>
<th>Structure</th>
<th>Origin and Uses</th>
<th>Citation(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Image of polyvinyl alcohol]</td>
<td>Polyvinyl alcohol is a major component of starch (including cornstarch)</td>
<td>Wikipedia (Amylopectin)</td>
</tr>
<tr>
<td>[Image of polyvinyl acetate]</td>
<td>Monomer unit is glucose thickener used in food</td>
<td>Kirk-Othmer Chemical Encyclopedia, <em>Starch</em> (Dec 4, 2000)</td>
</tr>
<tr>
<td>[Image of guar gum]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[Image of amylopectin]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3. Go to the Elmer’s glue website at [www.elmers.com](http://www.elmers.com). Locate the MSDS for Elmer’s Glue-All. What information is available about the contents of Elmer’s Glue-All?
4. Vinyl alcohol polymerization and crosslinking with borate

![Polyvinyl Alcohol](image)

a. What is the molecular formula of the repeating unit in polyvinyl alcohol?

b. Draw a borate ion that connects the two different PVA chains with 4 hydrogen bonds. Use dotted lines to show the interaction between the hydrogen bond donors and the hydrogen bond acceptors (there are many correct ways to complete this task).

c. For one of the hydrogen bonding interactions you have drawn, label the bond dipole, hydrogen bond donor, and hydrogen bond acceptor.

5. The number and type of intermolecular interactions has a direct effect on the macroscopic (observable) properties of a material. For example, consider the macroscopic physical property of viscosity.

a. Define viscosity.

b. How do you expect the viscosity of an aqueous solution of polyvinyl alcohol to be affected by addition of a crosslinking agent such as borate ions? (Think about the main intermolecular interaction between PVA and water. How will the borate ions affect that interaction?)

c. How do you expect the viscosity of this solution to be affected by the concentration of borate?

d. How do you expect the viscosity of this solution to be affected by the concentration of polyvinyl alcohol?

6. Read through the procedure. In the table in part 2 of the experiment, finish filling in the remaining values in the “Mass Ratio PVA:Guar Gum” column. Use the example in the introduction to help you with those calculations.

7. Fill in the table below by identifying the variables involved in the experiments you conduct in Part 1. Some example answers are given, but they may not be complete. The methodological and outcome variables are the same for each sample, so only one box is given for each of these categories.
Variables in sample preparation: Amount of PVA, amount of Borax, amount of water

Methodological variables: Stirring

Outcome variables: Stickiness

8. Where should you put the waste for this experiment? Read through the experiment and complete the following table.

<table>
<thead>
<tr>
<th>Item</th>
<th>Blue Chemically Contaminated Waste Bins</th>
<th>Regular Trash</th>
<th>White One Gallon Waste Bucket</th>
<th>Cleaned thoroughly and put back in the locker</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solutions of polymers</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Solid polymer samples</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Used reaction cups</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Used pipettes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Used Glassware</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Complete this experiment with a partner. To save time, split up the sample preparation work and experimentation with your partner, so that you are each working on a separate sample simultaneously. Make sure you both record all of the observations.

For Parts 1 and 2:

Obtain ~10 ml of each of the following aqueous solutions: 1% cornstarch, 50% Elmer’s glue, and 4% borax. Obtain ~17 mL of 0.5% guar gum and 4% polyvinyl alcohol (PVA). Use a clean glass beaker for each solution. Shake the stock bottle of guar gum and cornstarch before transferring to beakers. **If these solutions are on stir plates, make sure to return them to the plates after you obtain your sample.**

Obtain ten plastic reaction cups and five plastic transfer pipets. You will use one pipet for each solution. Reaction cups can be cleaned and reused if more trials are needed.

**Part 1: Exploration of Polymer Properties**

**Goal**

In this part of the experiment, you will prepare five cross-linked polymer samples and explore their properties.

**Procedure**

Label five of the reaction cups with letters A-E. Place 4 mL of a polymer solution in each reaction cup, according to the formulation instructions below.

A. 4 mL of 4% polyvinyl alcohol (PVA)
B. 4 mL of 1% cornstarch
C. 4 mL of 0.5% guar gum
D. 4 mL of 50% Elmer’s glue
E. 3 mL of 4% PVA and 1 mL of 0.5% guar gum.

Observe each solution in the sample cup. Record your observations.

Is each solution cloudy or clear? Is there any sediment in the bottom of the cup? What else is notable about each solution?

Add 1 mL of borax to each vial. Swirl gently to mix.

What happens?

Stir the sample vigorously with a glass stirring rod.

Does stirring cause a change? Do further changes occur after resting for a few min?
Squeeze a small amount of sample between your fingers.

Does it stick to your gloves?

Place a small amount of sample on a watch glass.

Does the sample stick to a watch glass?

Form the sample into a sphere. Hold opposite sides and slowly pull apart.

Does it hold its shape after stretching? Does it matter if you stretch the sample quickly or slowly? How far does it stretch before breaking?

Does it bounce when dropped from 12” onto the lab bench? If so, how high?

Place your sample back in the vial. Add a few more drops of Borax solution and swirl the vial to mix it.

What happens?

Conclusions

Consider all of your observations. With your group, discuss how the properties you have observed above relate to each other (property-property relationships) and how they relate to the structures of the cross-linked materials (structure-property relationships).

Your GSI will lead a class discussion. Take notes on the conclusions from this discussion.

Part 2: Exploration of Polymer Mixture Ratios on Properties

Goal

In this part of the experiment, you will explore the effect of mixing guar gum and PVA on the maximum length that your sample can stretch.

<table>
<thead>
<tr>
<th>Borax (mL)</th>
<th>PVA (mL)</th>
<th>Guar (mL)</th>
<th>Total volume (mL)</th>
<th>Mass Ratio PVA:Guar</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>1</td>
<td></td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>1</td>
<td></td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>1</td>
<td></td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>1</td>
<td></td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>J</td>
<td>1</td>
<td></td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

What is the model being tested here? Hint: it involves the crosslinking differences between PVA and guar gum and how they might impact stretchiness. Write a hypothesis for this experiment (which includes five trials, F through J).
**Procedure**

Devise a procedure to quantitatively determine the maximum stretch length of a crosslinked polymer sample.

Describe this method of measurement.

Label the remaining five reaction cups F-J. Prepare the samples (F-J) described above. Vigorously stir each sample with a glass rod before analyzing the maximum stretching length according to the procedure you designed.

Graph the results of your investigation. Is the graph linear? Does it have a maximum?

Describe any challenges you encountered while measuring the maximum stretch length of your polymer mixtures.

**Conclusions**

Your GSI will lead a discussion on property/property relationships. Summarize this information.

**Cleanup:** All polymer solutions should be poured into the appropriately labeled white, 1 gallon waste bucket provided in the hood. Double check that the bucket is lined with a plastic bag.

Empty reaction cups and used plastic pipettes can be disposed of in the blue “chemically contaminated materials” bins. Any glassware that you have used (such as watch glasses, beakers, and glass stirring rods) should be first rinsed into white, 1-gallon waste bucket with dH2O bottle, then washed with soap/Cleaning solution, rinsed with tap water, and then rinsed with distilled water. Verify that you have returned these clean items to the correct location in your lab drawer before you leave.
This page is intentionally left blank.
Name ____________________________________________

Partner(s) ____________________________________________

**Part 1: Exploration of Polymer Properties**

**Table 1:** Properties of polymer samples mixed with borate crosslinker.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Polymer solution (4 mL per vial)</strong></td>
<td>4% PVA</td>
<td>1% cornstarch</td>
<td>0.5% guar gum</td>
<td>50% Elmer’s glue</td>
<td>PVA (3mL) and guar gum (1 mL)</td>
</tr>
<tr>
<td>Observations before adding 1 mL of borax solution.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations after adding borax</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did stirring cause a change?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Do further changes occur after resting for a few minutes?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Does it stick to your gloves?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Does it stick to a watch glass?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Does it hold its shape after stretching?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Does it matter if you stretch the sample quickly or slowly?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How far does it stretch before breaking?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Does it bounce? If so, how high?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
9. Using the data in your table, try to explain some trends you see in the following polymer properties based on the structures of the polymers (i.e. formulate some structure-property relationships).

   A. Stickiness:

   B. Viscosity:

   C. Bounciness:

**Part 2: Exploration of Polymer Mixture Ratios on Properties**

10. Describe the model being tested in Part 2. Make sure you draw on your understanding of molecular structure in constructing your model.

11. What’s the hypothesis for this experiment?

12. Describe the method you used for determining the maximum stretch length of the polymer. How does this method produce the most robust stretch length data, as compared to other methods (if your method proved problematic, explain how it could be improved)?
**Table 2**: Effect of mixing Guar gum with PVA on the stretching properties of the samples.

<table>
<thead>
<tr>
<th>Borax (mL)</th>
<th>PVA (mL)</th>
<th>Guar (mL)</th>
<th>Total volume (mL)</th>
<th>Mass Ratio PVA: Guar</th>
<th>Stretch Length (cm)</th>
<th>Other Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>1</td>
<td>4</td>
<td>0</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>J</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Graph 1**: Effect of mixing Guar gum with PVA on the stretching properties of the samples.
13. Is the graph linear? Does it have a maximum?

14. Describe any challenges you encountered while measuring the maximum stretch length of your polymer mixtures.
15. After reviewing the results of your experiment…
   a. …what is the predominate trend that you see?

   b. Can you explain this trend using explanations based on molecular structures of the compounds involved?

   c. Based on your answer to part b of this question, what changes should you make to your initial model? (Note: if the answer is “none,” you must justify this.)
16. You have been promoted to Chief Sustainability Officer for the toy company. Your R&D division has given you three possible starting materials for making a toy, and you have to choose which one to use. From doing this week’s lab, you have some information about the properties of these polymers that would make a fun toy. However, as the CSO you also need to consider effects on human health and the environment. Based on your experience of doing the lab and the data shown in the table below, choose a polymer and make an argument as to why you chose it. Note: there are no wrong answers.

<table>
<thead>
<tr>
<th>Polymersic Starting Material</th>
<th>CAS No.</th>
<th>Human Health Effects</th>
<th>Persistent in the Environment</th>
<th>Accumulates in Organisms</th>
<th>Fish LC50 (µg/L water)</th>
<th>Soluble in Water</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guar Gum</td>
<td>9000-30-0</td>
<td>none known</td>
<td>unknown</td>
<td>unknown</td>
<td>218,000</td>
<td>yes</td>
</tr>
<tr>
<td>Polyvinyl alcohol</td>
<td>9002-89-5</td>
<td>suspected</td>
<td>no</td>
<td>no</td>
<td>86,000-118,000</td>
<td>yes</td>
</tr>
<tr>
<td>Polyvinyl acetate</td>
<td>9003-20-7</td>
<td>suspected</td>
<td>yes</td>
<td>no</td>
<td>data not available</td>
<td>no</td>
</tr>
</tbody>
</table>

Data retrieved from EPA databases: ACToR (http://actor.epa.gov) and ECOTOX (http://www.epa.gov/ecotox/)
Polymer Properties and Applications
Part 2: Cross-Linking Reactions in Toy Design

Introduction

In this week’s lab experiment you will continue to build on the same procedure and materials you used last week (sodium borate to cross-link solutions of guar gum, polyvinyl alcohol, polyvinyl acetate, and corn starch) to build a ‘toy’ and optimize one property of that toy. You will then enter your product in contest against other toys with the same optimized property. You will continue to explore the structure and behavior of polymer crosslinking and how the chemical structure affects the macroscopic (observable) properties of the materials. This experiment allows you practice designing your own experimental procedure in the context of product design and product development.

Research Questions

How can we apply what we have already learned about polymers to a specific application? How can we optimize the properties of a polymer to create a fun toy for children?

Science in Practice: Last lab, you performed experiments exploring polymer crosslinking to prepare you for this week’s lab, during which you’ll be creating a fun toy based on a particular polymer property. To make the best toy possible, you’ll need to optimize your property (i.e. make your polymer the stickiest or the stretchiest or the bounciest it can be). This means you’ll have to create a model for how your property works, develop experiments to test and refine your model, and ultimately share your methods with fellow classmates as you engage in a competition to determine who has made the best toy. You’ll be putting all that you’ve learned about models and experiments in the past weeks to work.

Recall the various experiments from Part 1 of the last lab. How many variables did you identify in each experiment in the pre-lab question? Can you figure out how many experiments you would need to do if you wished to alter one variable at a time to test all possible configurations of your model? Answer: at least $2^n$ (assuming one experiment per variable), where n is the number of variables. That can quickly turn into quite a lot of experiments. This is another example of the usefulness of models: by positing an explanation of how your phenomenon works, you can pick out the most useful experiments to execute in order to refine your model, reducing the time and material you’d need for doing those $2^n$ experiments.

When you start this lab, think critically about the most useful experiments to do. Try not to simply pick the first three experiments you can think of; be judicious about how much time you have in a 3-hour lab period to come up with the most robust model possible. Then, consider how best to communicate your results to your lab mates.
Understanding how scientists communicate with each other, and why, is a vital part of “doing science”. Science isn’t something that can be accomplished individually; scientists work together at many levels to ensure a rigorous standard of knowledge is upheld. It happens informally in a research group setting or in partnering with other lab groups (as you’re doing today). It also happens more formally at scientific conferences and with peer review of published scientific work. Sharing experimental procedures and data is a cornerstone of scientific practice. It helps ensure that conclusions being made from the models being developed (which could impact governmental policy, health practices, and many other things) are valid. You’ll get to experience this first hand this week as you and your classmates agree on how to judge the toys you’ll be developing and ultimately determine a winner of the competition.

After doing this lab, you should have a better understanding of these ideas:

**Scientific Inquiry:** How can we formulate the best set of experiments for refining a model?

**Science as a Social Practice:** How do scientists collaborate to determine what the most valid model is?
17. Define the terms dependent variable and independent variable. Look back at part 2 of last week’s experiment. What was the dependent variable and what was the independent variable? Explain.

18. Define the terms quantitative measurement and qualitative measurement. Give an example of each type of measurement from the data you collected last week.

19. Imagine you are working for a toy company and you want to create a polymer that would be fun for a 6-year old to play with. Consider the items below, which will help you to design and measure your results efficiently.

   a. List five properties of cross-linked polymer mixtures that you could optimize to create a fun toy.
   b. For each property, describe how you would quantitatively measure it.
   c. List three independent variables that you could explore when creating your samples this week.

20. In this lab, you will be designing your own methods for testing and optimizing a specific property of a polymer. Why is it important to change only one variable at a time throughout your tests? What would happen if you changed two variables simultaneously?

21. Where should you put the waste for this experiment? Read the experiment and complete the following table.

<table>
<thead>
<tr>
<th>Item</th>
<th>Blue Chemically Contaminated Waste Bins</th>
<th>Regular Trash</th>
<th>White One Gallon Waste Bucket</th>
<th>Cleaned thoroughly and put back in the locker</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solutions of polymers</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Solid polymer samples</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Used reaction cups</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Used pipettes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Used Glassware</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Complete this experiment with a partner. To save time, split up the sample preparation work and experimentation with your partner, so that you are each making separate sample simultaneously.

**Part 1: Selection of Contest Parameters**

**Goal**

You will brainstorm properties for a fun toy with your class and decided how to measure these properties for the toy competition later in the experiment.

**Procedure**

Your GSI will compile all of your ideas for an ideal property for a fun toy on the board. Each class will choose four properties that would make a fun toy polymer, and you will have class competitions to create polymers that optimize each property. With your partner, choose which of the properties you would like to optimize. Your goal will be to create a formulation for a polymer solution that optimizes one of these properties.

Your GSI will then divide your class into four groups - one group for each property.

> With this group, decide how to measure the property in the competition and how the samples will be judged. This must be quantitative.

After each group has discussed the competition judging, the resulting “contest rules” for judging each of the four properties should be shared with the rest of the class.

**Part 2: Optimization**

**Goal**

You will work with your partner to optimize one of the four properties chosen by the class in the previous part of the experiment.

**Procedure**

With your partner, identify which one of the four properties chosen by the class in part 1 you would like to optimize. Choose three different variables that you can change and plan experiments to test the effects of these variables on your chosen property. Keep the following guidelines in mind:

1. Each sample is limited to a total volume of 5 mL.
2. There may be additional materials located in the hood that you may use.
3. For each sample, you and your partner should independently measure the desired property and record your data. Try to perfect your technique to decrease the difference between your two values.

To plan your experiments, begin with an initial model concerning your property. You will plan a series of experiments with the ultimate goal of optimizing for the desired toy properties. Develop three hypotheses, and write them above the graphs where you’ll record your data for each experiment. Record your experiments in both table and graph format. **Your first three graphs must each have a different independent variable** from each other.

   Include at least four data points (four samples) per table/graph. Your results and your partner’s results should be listed in separate columns in the table.

   The independent variable (graphed on the x-axis) should be the variable you are currently testing. The dependent variable (graphed on the y-axis) should be the property you are optimizing, quantitatively measured according the competition rules.

Based on your results from your first three sets of experiments, refine your model. Plan and carry out at least two additional sets of experiments.

   Record the results of these experiments. These graphs may have the same independent variable as each other (and as prior experiments) if you wish.

When you and your partner have finished testing your hypotheses and have found a formulation that optimizes the desired property, write out clear instructions for making the best sample.

   Record these “formulation instructions”.

Part 3: Contest

Goal

You will write manufacturing instructions for another lab group. They will follow your instructions to build your toy and enter it into the toy contest. There will be one winner per property. The writer of the manufacturing instructions wins, not the manufacturer.

Procedure

When instructed by your GSI, give your formulation instructions to the “manufacturing department” (trade with another lab group). Follow the instructions provided by the other group to make a sample of their optimized crosslinked mixture for the contest.

Follow your GSI’s instructions for competition judging. May the best toy win!

Cleanup: All polymer solutions should be poured into the appropriately labeled white, 1-gallon waste bucket provided in the hood. Double check that the bucket is lined with a plastic bag.

Empty reaction cups and used plastic pipettes can be disposed of in the blue “chemically contaminated materials” bins.

Any glassware that you have used (such as watch glasses, beakers, and glass stirring rods) should be first rinsed into white, 1-gallon waste bucket with the bottle of distilled water, then washed with soap/cleaning solution, rinsed with tap water, and then rinsed with distilled water. Verify that you have returned these clean items to the correct location in your lab drawer before you leave.
Report Sheet

Name__________________________________
Partner(s) ______________________________

Part 1: Selection of Contest Parameters

1. List the four competition categories, and a brief description of the quantitative method designed by your class for testing each of these four properties

A.

B.

C.

D.

Part 2: Optimization

2. Which property will you and your partner optimize?

3. What is your initial model for your property? Identify three independent variables in your model that you can test to determine their effects on the property of interest. Hint: the data you gathered in the last experiment in Part 1 will help you with this.

Model 1:

Independent variables:
Hypothesis 1: ________________________________________________

Table 1

Graph 1

Conclusion from experiment 1.

How does this conclusion inform your choice for the next experiment?

4. Hypothesis 2: ________________________________________________

Table 2

Graph 2

Conclusion from experiment 2

How does this conclusion inform your choice for the next experiment?
5. **Hypothesis 3:**

   

   Table 3

   Graph 3

   Conclusion from experiment 3

   How does this conclusion inform your choice for the next experiment?

   Based on the data you’ve gathered from your three experiments, write a final model for your property here. Briefly recount the process you underwent to reach this model.

   **Model 2:**

   

   **Process:**

   

   6. Record the instructions here that you will give to the “manufacturing department” for the synthesis of your optimized polymer formulation. Also, write these instructions out on a separate piece of paper to hand to them.
Part 3: Contest

7. Write observations from competition judging here. Who had the best model? Now that you’ve shared your data and methods with others, what should be the next step in optimizing your property?

Reflecting on the results:

8. How did intermolecular forces (IMFs) or crosslinking relate to your design? What decisions did you make that would optimize the properties? State the desired strength of IMFs (relatively) and how you chemically achieved these goals. Be sure to mention your decision in formulations and how you adjusted your formulations to achieve your goals.

9. Consider the different factors involved in judging the toy competition. How did your group come to a consensus on which toy was the best? Do you think this process reflects the methods used by the larger scientific community as they validate scientific knowledge? Why or why not?
Appendix 2. HLM Assumptions

Figure 1. Level 1 residual analysis for the two-level modeling of P2M2, as described in Chapter 5. The mixed command was used with the reml option due to the small number of groups.
Figure 2. Level 2 residual analysis for the two-level modeling of P2M2, as described in Chapter 5. The mixed command was used with the reml option due to the small number of groups.

The figures above resulted from an analysis to determine the validity of the assumptions made about the HLM analysis conducted in Chapter 5. The assumptions are that the distribution of
the error terms was normal, mean-centered, and uncorrelated. These assumptions are largely warranted, as can be seen from the figures of the histograms and qnorm plots above. The histograms include a normal curve for comparison. Making the same types of graphs for the HLM analysis of the other variables in Chapter 5 yielded very similar results, so only those in the above figures were included for simplicity.