

A Learning Goals Driven Design Model for Developing Science Curriculum*

Joseph S. Krajcik and Katherine L. McNeill

Center for Highly Interactive Classrooms, Curriculum and Computers for Education

School of Education

University of Michigan

&

Brian Reiser

Learn Sciences

School of Education and Public Policy

Northwestern University

*Authorship is in alphabetical order with all authors contributing equally to the conceptualization of the paper

Current reform efforts in science education strive to develop materials that align with local, state, and national standards such as those articulated by the American Association for the Advancement of Science (AAAS, 1993) and the National Research Council (NRC, 1996). Alignment with standards is critical for students to achieve on distal measures, such as state mandated tests (National Research Council, 2006). Although this alignment is currently a national goal, few, if any curriculum materials succeed in this endeavor. Moreover, even fewer models exist that can guide researchers and developers in how to develop such aligned materials.

Project 2061's review of middle school curriculum materials concluded that none of the nine middle school programs they examined were likely to result in the attainment of the standards, the key learning goals used in the analysis (Kesidou and Roseman, 2002). Their critique included that the materials covered many topics at a superficial level and focused on technical vocabulary. Moreover, the materials did not take advantage of what we know about student learning. For example, they failed to take into account students' prior knowledge, lacked coherent explanations of real-world phenomena, and did not provide students with opportunities to develop explanations of phenomena (Kesidou and Roseman, 2002). Kesidou and Roseman (2002) propose that new middle school science materials that reflect findings from learning research and that focus on key learning goals need to be developed that support teachers in promoting students learning of the key ideas in science. In order to address this need, we are currently designing middle school materials to align with national standards that take into account current findings in research on the teaching, learning and assessment of science. We refer to our materials as *Investigating and Questioning Our World Through Science and Technology* (IQWST). The goal of our work in IQWST is to use what is currently known about the teaching and learning of science to develop coordinated grades 6 – 8 middle school science curriculum and study the affect of those materials on student learning. One goal of IQWST is to develop a design model that will result in alignment with standards with student and teaching materials and assessments as well as curriculum materials that support students in meeting these standards.

How effective an assessment is depends on how well it aligns with materials and instruction to reinforce common learning goals (Pellegrino, Chudowsky, & Glaser, 2001). Aligned assessment is for determining the extent to which students achieve learning goals. The closer student assessment is aligned with curriculum and classroom practice, the more likely assessment data will provide an accurate picture of learning (Ruiz-Primo, Shavelson, Hamilton, & Klein, 2002). Assessments that closely align with an enacted curriculum and its learning goals may be more immediately usable by teachers and researchers for getting feedback about whether students are achieving goals and for adjusting curriculum and instruction accordingly. However, there are few models that can guide developers in how to produce aligned assessments, materials for teachers and students and assessments.

In this paper, we describe our design process for creating assessments that align with learning goals, curriculum materials and assessments. A central goal of our research work is to narrow the gap between assessment, materials, and learning goals through the process of learning-goals-driven design. We describe our design process for developing standards-based curriculum materials for use in middle schools and report on our first and second rounds of pilot testing one of our units. Our aim in this article is to illustrate how our design process and

multiple iterations of revision can result in materials that align learning goals, student and teacher materials, and assessments to promote student learning.

We first begin by describing of our theoretical perspective and assumptions derived from current research on learning, instruction and assessment. Next, we describe our initial design process using a *learning goals driven* design model, similar to a backward design model (Wiggins & McTighe, 1998) to develop an inquiry-oriented chemistry unit that supports in-depth understanding of chemistry concepts and scientific inquiry practices specified in national science education standards. Next, we report on how we used multiple sources of data from our first enactment to identify concerns within the curriculum materials and subsequent revision of the materials to better align the learning goals, materials and assessments. We then report on our second enactment of the curriculum materials and present findings. We hope not to simply show the value of these curriculum materials, but rather describe how using a learning goals driven design model and multiple data sources to revise materials can help to align learning goals, materials, and assessments.

Theoretical Perspective on Learning and Instruction

In the last two decades, learning scientists, educational researchers and cognitive scientists have articulated principles on how children learn science (Bransford, Brown, & Cocking, 1999; Donovan and Bransford, 2006). We use these principles of learning to design new curriculum materials for students and teachers. In designing curriculum materials for teachers and students, we build upon six major ideas from the literature: 1) active construction, 2) situated learning, 3) social interactions, 4) cognitive tools, 5) the structure of expert knowledge and 6) science as a way of knowing. Extensive discussion of these ideas are discussed elsewhere, (Krajcik, Blumenfeld, Marx, & Soloway, 2000). Singer, et al. 2000). Because of the focus of this paper on learning goals, we discuss the structure of expert knowledge and science as a way of knowing. These two ideas influence our assessment framework.

Structure of Expert Knowledge

Science teaching and assessment should focus on what we value the most in science. As stressed in *How People Know* (Bransford, Brown & Cocking, 2000), superficial coverage of numerous topics in a subject area needs to be replaced with in-depth coverage of few central ideas that allow key concepts, principles and ideas of that discipline to be understood.

The knowledge that scientists hold is developed around conceptual organizes or schemas that guide how they solve problems and make observations (Glasser & Chi, 1988). Scientific knowledge is characterized as hierarchical, highly organized with many connections and interrelationships between ideas. Scientists have their knowledge organized around core concepts and principles or “big ideas” about the discipline with connections between the various ideas. This conceptual structure allows scientists to apply their understandings fluently to solve problems, interpret new information and build new understandings (Bransford, et al., 2000).

To help students learn and make use of their understanding, we need to help them build conceptual frameworks like experts. Developing curriculum materials around the “big” or enduring ideas of the discipline of science (Smith, Wiser, Anderson & Krajcik, in press; Wiggins & McTighe, 1998) is one way to support students in developing large, overarching conceptual frameworks similar to those of scientist. Enduring ideas are essential in understanding the scientific discipline, explaining phenomena and at making connections to the personal and social

lives of the learners. In aligning curriculum materials with these “big ideas” of a discipline, curriculum designers can sequence materials in a clear and logical way to help students move from inert to useable knowledge.

In designing the IQWST materials, we used national standards to select our big ideas. However, because standards are often short declarative statements of knowledge claims, we found that we needed to expand on their meaning. In IQWST we unpack the standard to get clarity of all the related ideas in the standard to capture the big ideas of the discipline.

Science As a Way of Knowing

Constructivist and situated views of learning also align with the current view of scientific knowledge as socially constructed knowledge (Duschl & Hamilton, 1998). Scientific knowledge is now seen as a model that the current scientific community agrees upon to explain phenomena (Kuhn, 1970).

This view of science as a social construction and learners as constructing their own knowledge in a particular context means that knowing science involves students developing a different way of thinking. Students need to be enculturated into these scientific ways of knowing. Driver et. al (1994) discuss the importance of this shift in their view of learning science.

It means that learning science involves being initiated into scientific ways of knowing... learning science thus involves being initiated into the ideas and practices of the scientific community and making these ideas and practices meaningful at the individual level. The role of the science educator is to mediate scientific knowledge for learners, to help them to make personal sense of the ways in which knowledge claims are generated and validated... (page 6)

Our curriculum design reflects this argument. Knowing science content is not memorizing facts, but rather being able to use content in different contexts and with different scientific inquiry practices, such as developing evidence based explanations of phenomena.

In IQWST, we develop *learning performances* (described below) as the key learning goals of the curriculum units that reflect the reasoning tasks we want students to be able to do with scientific knowledge. Learning performances reformulate a scientific content standard in terms of scientific practices that *use* that content, such as students being able to define terms, describe phenomena, use models to explain patterns in data, construct scientific explanations, or test hypotheses. This step is necessary because standards are written in terms of declarative statements that do not specify what the student should do with the knowledge.

Learning Goals Driven Design Model

We apply the principles described above using a learning goals driven design model (see Figure 1). This model is a modification and expansion of backwards design (Wiggins and McTighe, 1998). The model includes three stages: 1) specifying learning goals, 2) materials development and 3) feedback. The learning goal stage has two steps: 1) identify and unpack national standards, and 2) develop learning performances to operationalize standards. The materials development stage has four steps: 1) contextualize the unit through a driving question and anchoring events, 2) identify learning tasks, 3) produce an instructional sequence, and 4) create assessments and rubrics all of which are linked to the learning performances. The

feedback stage includes 1) pilot test materials, and 2) receive feedback from external reviewers. Although these steps are listed linearly, the diagram in Figure 1 reflects the iterative nature of the process where the development of multiple steps could occur simultaneously and the later components of the design cycle, such as the assessments, informed previous steps, such as the learning performances. Each of these steps is described in more detail below.

We describe our design model focusing on aspects of assessment and illustrate the process by using examples from one of the first units developed in IQWST, How can I make new stuff from old stuff? (Stuff) (McNeill & Krajcik, in press). This unit focuses on chemistry content as our key learning goals. Science educators have long been concerned that many middle school students have difficulty learning basic chemistry concepts. Four central ideas in chemistry—the particle nature of matter, the conservation of matter, substances and their properties, and chemical reactions—are important for middle school students to grasp because they serve as a basis for learning other, more complex science ideas in physics, biology, and chemistry taught at the secondary level (AAAS, 2001). Yet, research on student learning has consistently shown that students have difficulty learning these concepts (Driver, Squires, Rushworth, & Wood-Robinson, 1994; Lee, Eichinger, Anderson, Berkheimer, & Blakeslee, 1993). We strove to design materials so that students can apply these chemistry ideas to explain a range of phenomena they observe in their everyday lives.

Learning Goals

Identifying and Unpacking Standards. We use the national science education standards to identify the big ideas we want students to learn. This step was informed by the importance of considering the structure of expert knowledge. We identify key learning goals from the nationally recommended science learning standards, *Benchmarks for Science Literacy* (AAAS, 1993), the companion document *Atlas of Science Literacy* (AAAS, 2001) and the National Science Education Standards (NRC, 1996). Using these documents, we created a concept map that included all the key content standards for middle school students that we would like students to learn in the unit as well as requisite prior knowledge and common misconceptions. These maps help us decide on the focus of the unit by showing which ideas link together. Figure 2 shows the concept map we developed.

Our next step was to “unpack” each of the standards. By “unpack” we mean that we broke apart and expanded the various concepts in the standard to elaborate the intended science content in the relatively succinct standards (Appendix A). For example, during the unpacking the content standard from The Atlas about chemical reactions (AAAS, 1990, p.47), we realized that for a student to understand this standard required scientific understanding of the terms “substance” and “property” and that these concepts needed further elaboration. We returned to the AAAS documents to find a standard related to substances and properties but there was not one listed. Consequently, we turned to the National Science Education Standards (National Research Council, 1996) and incorporated their middle school chemistry standard about “property” and “substance” (Appendix A). The unpacking not only resulted in the elaboration of particular learning goals, but also in adding a new learning goal to the unit.

Developing Learning Performances. Once we identified and unpacked the key content learning goals, we developed “learning performances” that require a range of cognition from students (Appendix B). This step was informed by science as a way of knowing. We derived our different ways of knowing for our learning performances from the recently revised Bloom's

Taxonomy (Anderson & Krathwohl, 2001) and from the habits of mind standards (AAAS, 1993) and the scientific inquiry standards (National Research Council, 1996). We develop learning performances by crossing the content standards various ways of knowing. Figure 3 illustrates the process of developing learning performances.

Figure 3: Developing Learning Performances

Content Standard	X	Scientific Practice Standard	=	Learning performance
When substances interact to form new substances, the elements composing them combine in new ways. In such recombinations, the properties of the new combinations may be very different from those of the old (AAAS, 1990, p.47).		Develop...explanation s... using evidence. (NRC, 1996, A: 1/4, 5-8) Think critically and logically to make the relationships between evidence and explanation. (NRC, 1996, A: 1/5, 5-8)		Students construct scientific explanations stating a claim whether a chemical reaction occurred, evidence in the form of properties, and reasoning that a chemical reaction is a process in which old substances interact to form new substances with different properties than the old substances.

We use these learning performances as learning outcomes instead of just using the standards because “knowing” science is more than just memorizing these succinct statements about science. We believe that to teach and assess students’ understanding of this content, the standards need to be operationalized into different learning performances. Our learning performances explicitly break down the singular generic concept of “knowing” into multiple ways of knowing by combining both the science content and scientific practices (Appendix B). Each learning performance addresses a different way of knowing the science content. A set of learning performances together, provides a more complete picture of a student’s understanding. These learning performance are central both in creating the instructional sequence and assessments.

Development Stage

After creating the initial learning performances, we develop the next four components: assessments and rubrics, learning tasks, instructional sequence, and contextualization. Although we iteratively worked on these steps with each one informing the others, below we discuss them in a linear manner. Because this manuscript focuses on assessment, we only discuss that version of development here. Because learning tasks are also used as assessments in our work, we include the development of learning tasks as well.

Learning Tasks. After unpacking the standards and developing learning performance, we created instruction tasks that have the potential to foster students developing understanding of the learning goals. This aspect of the design model was informed by the ideas surrounding active construction, social interactions and cognitive tools. When developing tasks, we started by identifying various phenomena that align with the learning goals and vividly illustrate the learning goal for the students. We strove to find phenomena that align with the learning goal, would make complex scientific ideas plausible to students and enhance students' sense of the

usefulness of scientific concepts (Kesidou & Roseman, 2004). For instance, one of the major learning goals for the unit is:

When substances interact to form new substances, the elements composing them combine in new ways. In such recombinations, the properties of the new combinations may be very different from those of the old (AAAS, 1990, p.47).

One phenomenon that aligns with this standard is making soap from lard and sodium hydroxide. Students engage in this first hand experience around this key science concept by working with their peers to both perform and make meaning of the phenomenon. Once various phenomena are identified, we used the learning performance to guide the development of various instructional tasks allowing the materials to become cognitive tools that helped structure student learning. For the learning performance:

Students construct scientific explanations stating a claim whether a chemical reaction occurred, evidence in the form of properties, and reasoning that a chemical reaction is a process in which old substances interact to form new substances with different properties than the old substances.

We developed a task in which students carried-out and observed the reaction to make soap and then used their observations to write a scientific explanation. Using learning performances as a guide to developing tasks helps to assure that the learning goals and instruction align.

Instructional Sequence. Once learning tasks were identified, we created an instructional sequence that proceeded in a logical manner to help build understanding and which provided information to answer the driving question. This sequence encouraged students' active construction of knowledge and allowed the materials to act as a cognitive tool for student learning. For instance, students explored and developed understanding of the following standard at the beginning of the unit:

A substance has characteristic properties, such as density, a boiling point, and solubility, all of which are independent of the amount of the sample (NRC, 1996, p.154)

We first developed understanding of this standard because it was necessary prior knowledge for the related standard on chemical reactions. The instructional sequence contained a number of investigations that students completed allowing them to cycle back to these ideas of substances and properties as well as build on them to develop an understanding of chemical reactions. Appendix C is the project calendar for the first enactment and describes the learning tasks involved in this sequence.

Assessments. We wrote assessment items that directly related to the learning performances. Appendix E shows the alignment among a learning performance and an assessment item. To guide the alignment process, we also developed rubrics to assess students' understanding of these learning performances (see Harris et al., in press and McNeill & Krajcik, in press for discussion of assessments and rubrics). The base rubrics correspond to the different cognitive processes articulated in our learning performances (e.g., define, identify, explain, design, analyze and

interpret). A base rubric articulates the different components of a particular way of knowing and the levels of those components. These base rubrics can be adapted to any science content and thus can be used across all science curricula. For instance, to assess student understanding of scientific explanation, we developed a base rubric to use across different content areas (Harris, et al. in press) (See Appendix D). We used our base rubrics to develop specific rubrics for assessing students on each learning and assessment task for our chemistry unit. Appendix D also includes the specific rubrics we used to score the two explanation tasks on the pre and posttest. The rubric includes the three components of scientific explanation (claim, evidence, and reasoning) and discusses the criteria for different levels of each component.

First Enactment: Methods, Results, Concerns, and Revision of Curriculum

Participants and Setting

Three teachers enacted the first four-week unit in three different locations, two urban areas and one rural area, during the 2001-2002 school year with a total of 209 seventh-grade students (see Table 1).

Table 1: Teachers, students, and classrooms involved in the first enactment

Site	2001 -2002 School Year		
	Urban A	Urban B	Total
Schools	1	1	2
Teachers	1	1	2
Classrooms	1	3	4
Students	31	88	119

The two urban sites, Urban A and Urban B, are large cities in the Midwest. Students from Urban A attended a public neighborhood school. The school is “typical” compared to other schools in the district. Most of the Urban A students come from lower to lower-middle income families (approximately half of the cities students live in families that are at or below the poverty line), are largely minorities (over 90% are African Americans), and are mobile. Student dropout rates are high and their test rates are low compared to other students in their state (Blumenfeld, Fishman, Krajcik, Marx, & Soloway, 2000). The students in this particular school were mostly African Americans and from lower to lower-middle income families. Students from Urban B also attended a public middle school. This student group was ethnically diverse, and from lower-middle to middle income families.

Data Sources

We collected a variety of data sources to measure student learning and to critique the strengths and weaknesses of the unit to help determine an alignment between learning goals, learning tasks and assessments. . The data sources included: student pre and posttests, student artifacts, field notes, selected classroom videos, teacher feedback, Project 2061 review, and content expert feedback. We used these different data sources to identify and triangulate our concerns about the unit that we then addressed in the subsequent curriculum revision.

Identical pre and posttest measures consisted of 20 multiple-choice and 4 open-ended items. For Urban A and Urban B, we only included students in the analysis who completed both the pre- and posttest assessments. Due to high absenteeism, only 12 students from Urban A took both pre- and posttest assessments and in the Urban B classes, 77 students completed both pre and posttest measures. We scored and tallied the multiple-choice responses for a maximum score of 20. We developed rubrics to score the four open-ended items with a maximum possible score of 15. Pairs of independent raters scored the open-ended items using appropriate specific rubrics with an average inter-rater reliability of 90%. A third independent rater resolved disagreements.

In the Urban A enactment, we collected and analyzed student artifacts. Similar to the open-ended test items, we used rubrics to score student artifacts. In this case, raters assigned scores through discussion and then came to agreement by consensus. We used the artifacts to confirm and disconfirm our results from the test analysis.

We also examined classroom field notes from Urban A classrooms looking for general themes across the entire curriculum enactment as well as specific incidences that appeared to represent strengths and weaknesses in the unit. Based on these themes and incidences, we watched selected videotapes from all three enactments looking for confirming and disconfirming evidence to support our hypotheses from the field notes.

Teacher feedback provided another data source on the quality and usability of the curriculum materials. We solicited teacher feedback through two different methods. First, all three teachers participated in weekly phone conversations where we asked them for their opinions of the curriculum as they were actually using the materials. We then held a wrap-up meeting after all three teachers enacted the unit to obtain their reflections looking back at the unit and to discuss more in depth the unit as a whole.

Finally, we received data from two external sources. Project 2061 performed a preliminary analysis using the criteria described in Kesidou and Roseman (2002). We examined the explicit written feedback we received from the Project 2061 to identify strengths and weaknesses. We also received advice from a content expert in chemistry when we had specific concerns about the content in the unit. For example, we had the expert read text or excerpts from the curriculum when we were concerned about the accurate portrayal of the chemistry content.¹

Overall Results

In order to create an overall picture of student learning during the curriculum we first discuss the general test results. Then we look closer at one of the key learning goals, chemical reactions, to discuss how we used the multiple data sources to identify a number of concerns and subsequently revise the unit.

Overall, we found that students achieved significant learning gains from the enactment of the unit. Table 2 contains the pretest and posttest performance data by site. Although the pretest scores suggest that the two groups of students began the unit with different prior knowledge, they both resulted in learning gains as illustrated by the effect sizes of 1.54 and 1.10. The analysis shows significant achievement gains, yet the scores on the posttest were still low on an absolute level. We consciously designed the pre and posttest measures to be difficult in order to challenge

¹ After the second enactment of the unit, we had two content experts complete a more thorough review of the curriculum materials where they read the unit in its entirety and evaluated the appropriateness of the content.

student thinking and to prevent a ceiling effect so we could more accurately tease a part students' strengths and weaknesses.

Table 2: Enactment 1 test data by site

Site	Pretest M (SD) ^a	Posttest M (SD)	t -Value ^b	Effect Size ^c
Urban A ($n = 12$)	10.29 (3.24)	15.29 (4.07)	4.09**	1.54
Urban B ($n = 77$)	14.73 (4.54)	19.73 (5.30)	11.13***	1.10

^a Maximum score = 35

^b One-tailed paired t -test

^c Effect Size: Calculated by dividing the difference between posttest and pretest mean scores by the pretest standard deviation.

** $p < .01$; *** $p < .001$

To establish a more precise image of students' strengths and weaknesses, we examined the tests results in terms of the two different content standards (Appendix B). We wanted to determine if there was a difference in student learning of substance and properties versus chemical reactions. Overall, students achieved significant gains for both standards (see Table 3).

Table 3: Enactment 1 test data for substances and properties versus chemical reactions

Site	Pretest M (SD) ^a	Posttest M (SD)	t -Value ^b	Effect Size ^c
Urban A ($n = 12$)				
Sub & Prop	4.00 (1.91)	6.92 (2.64)	4.61***	1.53
Chem Rxn	4.13 (2.05)	6.46 (2.48)	2.61*	1.14
Urban B ($n = 77$)				
Sub & Prop	6.32 (2.46)	9.27 (3.09)	9.52***	1.20
Chem Rxn	6.07 (2.24)	7.64 (2.20)	6.56***	0.70

^a Maximum score: Substance and Property = 15, Chemical Reactions = 16

^b One-tailed paired t -test

^c Effect Size: Calculated by dividing the difference between posttest and pretest mean scores by the pretest standard deviation.

- $p < .05$; ** $p < .01$; *** $p < .001$

Yet, the effect sizes indicate that students at both locations had greater learning gains for the substance and property content than the chemical reaction content. Although the learning gains occurred, we wanted to revise the unit in order to support greater student learning, particularly for the chemical reaction content. To inform our revisions, we further analyzed the test results as well as examined data from the other sources to identify specific concerns and possible revision solutions.

Concerns and Revision of Curriculum

Between the first and second enactment, we redesigned and extended the unit to eight weeks to address all five of the identified standards (Appendix A) about substance & properties, chemical reactions, the conservation of mass, the particle nature of matter and the conservation of mass in terms of the particle nature of matter. We focus our discussion here on one section of the unit to provide a more in depth picture of how we used multiple data sources to identify issues and revise the curriculum. We chose the section of the unit that focuses on chemical reactions because of the lower gains we observed for this learning goal in the first enactment. However, we used a similar process for the revision of the other aspects of the unit.

Our analysis resulted in eight concerns and subsequent changes in the chemical reaction portion of the curriculum unit. Table 4 presents a summary of each concern, data sources, and evidence from the data source. We specifically discuss each of these in greater detail as well as explain the changes we made to the unit to address each concern.

Concern #1: alignment of learning performances and standards. Project 2061's review found that our learning performances were not sufficiently aligned with the learning goals, which they identified as the national standards. Furthermore, they argued that the learning performances could not substitute as "knowing" the standards (see Table 4). We were concerned about the alignment of our learning goals and that our rationale for using learning performances did not seem to be explicit in the unit.

Although we realize that student performances are not "knowledge", they provide us with an indication of whether students have acquired the understanding related to the performance. Our assumptions about learning performances stem from the theoretical perspective we presented earlier. In order to address Project 2061's concern, we rewrote the learning performances to clarify the language for better alignment with the standards as well as to make our assumptions about learning and instructions explicit in the learning performances. In order to integrate our assumptions about the importance of inquiry abilities as a scientific way of knowing, we added inquiry ability science standards (NRC, 1996) and habits of mind standards (AAAS, 1993). We specifically linked each learning performance to both content and inquiry standards to refine and articulate the different ways of knowing the science standards (Appendix F). We carefully used the same ways of knowing, such as creating scientific explanations, designing experiments, and using models, across the different content standards. For example, we had students write scientific explanations for whether two substances are the same, whether a chemical reaction occurred, and whether mass changed in order to assess students understanding of the content as well as their ability to write scientific explanations. We carefully chose our language and explicitly connected each learning performance to the appropriate science and inquiry standards. Furthermore, we revised the language in the instructional materials in order to make the connections between the lessons and the standards more explicit.

Concern #2: students' construction of scientific explanations. Project 2061's review of our learning performances, specifically articulated a concern that the learning performances had not clearly defined what is meant by "explain" (see Table 4). We developed an instructional framework for scientific explanation in which we broke the practice of scientific explanation into three components: claim, evidence, and reasoning (McNeill & Krajcik, in press; Moje et al. 2004). Yet in looking back at the first version of the curriculum we realized these criteria were not explicit in either the student or teacher materials. When we analyzed students' pre and posttests written explanations, we also had some concerns around student learning of scientific explanations (Table 5).

Students in Urban A did not achieve significant learning gains for explanation as a whole or for any of the individual components. Although students in Urban B did have significant learning gains for explanations as a whole including the evidence and reasoning components, the effect sizes were lower than we would have preferred. Looking at students mean posttest scores for evidence and reasoning at both sites shows that students had difficulty with these aspects of constructing scientific explanations particularly with reasoning.

Table 4: Concerns, Data Source, and Evidence for Chemical Reaction Section of the Unit

Concern	Data Source	Evidence from Data Source
#1 Learning Performances were not sufficiently aligned with standards	Project 2061	“While the learning performances can serve to operationalize what students might be expected to do with the knowledge in the learning goals, the performances do not substitute for the knowledge” (Review #1, page 2). “...there is not a good match between learning goals and learning performances...it’s important to check the performances against the learning goals to be sure that a) the knowledge in the learning goal is needed for the performance and b) students can carry out the performance with the knowledge in the learning goal” (Review #2, page 2).
#2 Curriculum materials did not explicitly state what is meant by “explain” and do not provide teachers or students guidance in creating “Explanations”	Pre and Posttests	Showed student difficulty in including the evidence and reasoning components of a scientific explanation (see table 5)
	Student Artifacts	Showed student difficulty in including reasoning component of a scientific explanation (Harris et. Al., in press).
	Project 2061	“...it is not clear what is meant by the phrase “explain that” in learning performances 3 and 9.” (Review #1, page 2)
#3 Student Readers did not include enough breadth and depth of content	Fieldnotes and videotapes	Teachers used the Student Readers to help students understand the concepts of substance, but there was not a similar portion of the reader for other concepts.
	Feedback from teachers	Teachers suggested expanding the reader because it provided opportunities for discussion, reflection, and helped create links between different concepts and activities.
#4 Student readers were not sufficiently integrated into the teacher materials or give teachers enough support in their use.	Feedback from teachers	Teachers articulated that the materials need to provide more guidance on how to use the reader.
	Fieldnotes and Videotape	Classroom enactment revealed that the reader was not being used in the method envisioned by the curriculum designers.
#5 Properties were not integrated into the chemical reaction portion of the unit	Project 2061 review	“students spend all this time up front investigating density, solubility, melting point as characteristic properties of substances but then do not use these properties in subsequent lessons to establish whether a chemical reaction has occurred or not. Instead, different properties (thickness/runniness or state of the material) turn up without explicit preparation.” (Review #1, page 7)
#6 After the unit, students thought mixtures were a chemical reaction	Pre and Posttests	On multiple-choice item 13, more students thought making lemonade was a chemical reaction after the unit than before. On the open-ended items, students included “dissolving” as evidence for a chemical reaction.
	Content expert	Discussed with content expert the distinctions between dissolving and chemical reactions as well as the most appropriate way to discuss this with middle school students.
#7 Atoms and molecules were not included in the unit	Project 2061 review	“Elements are not mentioned...If the idea of a small number of elements is not to be developed until later, what is the rationale for this approach? Four weeks seems a large investment for a small payoff.” (Review #1, page 6)
#8 Pre and Posttest did not adequately align with learning goals.	Project 2061 review	“It is of some concern that the Test at the beginning is not well focused on assessing the learning goals or key ideas...For nearly all of the items, the key ideas are either not necessary or not sufficient for responding correctly.” (Review #1, page 10)
	Pre and Posttests	Suggested that some of the questions did not effectively assess the desired learning goal.

Table 5: Enactment 1 data for scientific explanation

Site	Pretest $M (SD)^a$	Posttest $M (SD)$	t -Value ^b	Effect Size ^c
Urban A ($n = 12$)				
Total	1.88 (1.75)	3.19 (1.75)	1.91	0.75
Claim	1.25 (1.31)	2.08 (0.97)	1.77	0.63
Evidence	0.52 (0.78)	1.01 (0.96)	1.77	0.63
Reasoning	0.10 (0.36)	0.10 (0.36)	0.00	0.00
Urban B ($n = 77$)				
Total	2.85 (1.49)	3.71 (1.25)	4.61***	0.58
Claim	1.95 (1.04)	2.21 (0.81)	1.92	0.25
Evidence	0.74 (0.72)	1.21 (0.83)	4.15***	0.65
Reasoning	0.16 (0.42)	0.29 (0.53)	2.04*	0.31

^a Maximum score: Total = 9, Claim = 3, Evidence = 3, Reasoning = 3

^b One-tailed paired t -test

^c Effect Size: Calculated by dividing the difference between posttest and pretest mean scores by the pretest standard deviation.

- $p < .05$; ** $p < .01$; *** $p < .001$

We then analyzed the explanation students wrote as part of the classroom instruction (See Harris, McNeill, Lizotte, Marx, & Krajcik, in press for more details). We looked at specific target students' explanation written over the course of the unit. From this analysis, we found that while students claim and evidence improved, they did not include reasoning in their explanations. For example, for one chemical reaction task we were looking for students to articulate that a chemical reaction is when substances interact to form new substances with the new substances having very different properties from the old substances. A typical students' reasoning included "This evidence supports that a chemical reaction occurred because you can follow the evidence and determine that it change." Students' reasoning rarely justified why their evidence supported the claim by using the underlying scientific principle.

In order to address these concerns, we made a number of changes to make scientific explanations more explicit in the unit and to provide both teachers and students support in accomplishing this complex task. We added scientific explanation standards to our learning goals (NRC, 1996) and revised the learning performances to include claim, evidence, and reasoning (Appendix F). We also added a lesson to the unit where teachers introduce scientific explanation to the students through a variety of instructional strategies including defining scientific explanation, modeling examples of explanations, and critiquing explanations. Furthermore, we added written curricular scaffolds to the student materials, which supported students with each of the components (McNeill, Lizotte, Krajcik & Marx, in press). Finally, we revised the pre and posttests to include items that assessed explanations across the different content standards in order to give us a more complete picture of student understanding.

Concern #3: breadth and depth of student readers. In their feedback, teachers requested that we increase the breadth and depth of the student readers, which initially only included reading materials for three of the eight lessons. For example during the wrap-up meeting for the curriculum unit, both teachers from Urban A and Urban B commented on the importance of the reader and that they would like to see it expanded. The Urban B teacher said that it provided more time for discussion and encouraged student reflection. The Urban A teacher agreed and

added that the reader also helped students make links between different laboratory investigations and the content.

We examined the fieldnotes and videotapes to investigate how the teachers and students used the readers. During the substance and properties portion of the unit, we found that teachers used the student reader to elicit students' prior understandings about properties and substances. For example, one question in the reader asked students to identify two objects that are made of the same substance. In the Urban A classroom, one group of students believed that a toothbrush and toothpaste were the same substance because they are both used for cleaning teeth. This led to a class discussion about the difference between what objects are used for versus what makes up the objects. The teacher used the question in the reader to make students thinking visible as well as to help clarify the scientific definition of substances.

In the revision of the unit, we expanded the student reader to address more of the concepts in the unit as well as to provide more opportunities for reflection and discussion. We also based these revisions on guidelines about constructing materials to promote literacy in the sciences (Moje, et. al, 2004). For example, because of the conversation about the toothbrush and toothpaste, when we revised the reader we added a section to help students distinguish between objects' use and their composition. Another addition that we included was based on the Project 2061 review that suggested that we incorporate the Rumpelstiltskin story into the unit. Before students complete their first chemical reaction, we added an abbreviated version of the Rumpelstiltskin fairy tale into the reader. After the students read the story, we included a question that asks students whether they think straw can be turned into gold. This question makes visible students' prior understandings about old substances turning into new substances. After students explore a chemical reaction in class, students then return to this question of straw to gold in the reader to help them reflect on their prior understanding.

Concern #4: providing support for using the student reader. During the wrap-up meeting, the teachers also requested more information on how to use the student reader. The teacher from Urban A suggested including more guidance in the teacher materials in how to use the student reader in general as well as for a jumping off place for discussions. The teacher from Urban B agreed with her request. In reviewing the fieldnotes and videotape, we found that the teachers predominantly used the reader as the focus of an entire class period (for example, doing round-robin readings of entire sections). We envisioned teachers using the reader in a variety of ways such as an introduction to an experiment, as a journal topic at the beginning of class, as a homework assessment, as a way to begin a whole class discussion, or as a formative assessment. While we imagined different possibilities, they were not explicit in the curriculum.

When we revised the curriculum, we integrated the reader into the teacher notes suggesting places in the lessons and a variety of ways to use the reader. For example, we suggested that an open-ended question in the reader could be answered for homework and then reviewed in class the following day; it could be used as an in-class writing activity (e.g. bell work) preceding a lesson, or it could be used as a summarizing activity to close a class period. In all cases, the students first encounter concepts in class and then use the reader to extend, apply, summarize, clarify, or contextualize in a different phenomena. We also created an annotated reader for the teachers, which further elaborated on different ways to use the reader as well as included possible student responses. The "sample" student responses included ideal responses, acceptable responses that might generate further discussion, and inappropriate responses that

might signal student misconceptions. In the latter case, we also provided strategies for how the teacher might address those misconceptions.

Concern #5: properties integration in chemical reaction section. In Project 2061's review, they critiqued our lack of integration of properties into the chemical reaction segment of the materials. In the beginning of the unit, students determined melting point, solubility, density, hardness, and color of both fat and soap. While they again investigated these properties for their homemade soap, they did not use melting point, solubility or density in the interim lessons. Project 2061 voiced their concern that students would not integrate their understanding of properties with chemical reactions because of this disconnect.

We were concerned about students determining the density, solubility, and melting point for the substances in all experiments both because of the time to complete these procedures and the difficulty of making measurements for some substances (e.g. the density of oxygen is 0.00131 g/cm³). Consequently, we decided to resolve this issue by revising some chemical reaction experiments so that students collected data on the properties of the substances while for other experiments we provided students with the properties of the substances in a table. For example, for the electrolysis of water experiment, students would not be able to determine the melting point, solubility, and density for the hydrogen and oxygen gas. Consequently, we provide students with a table with this information. Then they discuss as a class what this information tells them and why they cannot determine the properties themselves.

Concern #6: students thought mixtures were a chemical reaction. In examining each item on the pre and posttest, we found a discouraging trend in the student data for one question. One multiple-choice question asked "Which change will produce a new substance?" Options were: a. Dissolve lemonade powder in water, b. Burning a candle, c. Heating water until it evaporates and d. Stretching a rubber band. Students' responses on the pre and posttest are in Table 6.

Table 6: Student responses to an item identifying that a new substance formed (n=89)

Possible Response	Percentage on Pretest	Percentage on Posttest
a. Dissolve lemonade powder in water	55.1%	67.4%
b. Burning a candle	15.7%	16.9%
c. Heating water until it evaporates	27%	14.6%
d. Stretching a rubber band	2.2%	1.1%

Although the correct response to this item is "b. Burning a Candle" the majority of students on the posttest select "a. Dissolving lemonade powder in water." This suggests the unit may have encouraged students' conceptions that mixing things together always results in a chemical reaction. In the open-ended responses, we also found that some students wrote that "powder dissolving" counted as evidence for a chemical reaction (Harris, McNeill, Lizotte, Marx, & Krajcik, in press). We discussed with a content expert the difference between dissolving and chemical reactions and the most appropriate way to discuss this with middle school students. Based on his suggestions, we modified the curriculum unit. We discuss these modifications below.

In order to address students' misunderstanding that mixtures are chemical reactions, we added one lesson to the unit specifically focused on mixtures. Students create a mixture and examine the properties before and after to determine if a new substance is made. Furthermore, they analyze particle models of chemical reactions, phase changes, and mixtures and discuss the

similarities and differences of these processes. We also added an explicit section of the student reader to address this concern as well as suggestions for teachers on how to lead discussions around these ideas in the annotated reader.

Concern #7: including the particle nature of matter. Project 2061's critique included that we did not cover all of the chemical reaction standard (AAAS, 1990, p.47) because we did not include the concepts of the particle nature of matter. Furthermore, they stated that this was a large time investment to spend on chemical reactions and not cover the particle model. When we specifically asked the teachers who had enacted the unit if they thought that the particle nature of matter could be brought in earlier, they thought it sounded like a good idea and might in fact increase students' understanding. Originally, our outline of the unit included four learning sets: substances and properties, chemical reactions, conservation of mass, and the particle nature of matter in terms of both chemical reactions and the conservation of mass.

In order to address this concern, we revised the unit to include three learning sets with a focus on substance and property, chemical reactions, and conservation of mass with the particle nature of matter integrated throughout each learning set. We revised the unit to introduce the particle model in Lesson 42. After Lesson 4, the unit continuously cycles back to the particle nature of matter with modeling activities, sections of the reader, and discussions after students first hand experiences of the various phenomena.

Concern #8: pre and posttest alignment with learning goals. In their initial review of the unit, Project 2061 critiqued a number of the assessment items because they found that the learning goals (standards) were not necessary and sufficient to complete the items. Later, they performed an extended analysis of a number of our assessment tasks using five questions: 1. What knowledge is needed to answer this task? 2. What idea is likely being assessed? 3. Is the knowledge needed to correctly respond to the task? 4. Is the knowledge enough by itself to correctly respond to the task or is additional knowledge needed? 5. Will the task likely be an effective probe of this knowledge? They found that a number of our assessment items were not aligned with the standards from this particular perspective. We also examined each of the questions on the pre and posttest to determine if they aligned with our revised learning performances, assessed student learning of content and inquiry abilities, and included appropriate distracters for the multiple choice questions.

Based on both our and Project 2061's analysis, we discarded over half of the items on the pre and posttests. The items that remained were then revised. For example, the question discussed in concern #6 was kept because it seemed to address a common student misconception about chemical reactions. But we discarded the choice about stretching a rubber band because so few students selected this choice both before and after the curriculum. We also took into account our addition of the inquiry ability standards and revised learning performances. For example, we specifically included open-ended question about constructing a scientific explanation for both substances and chemical reactions so we could examine this way of knowing across different science content.

² We assumed that students had a basic understanding of the particle model of matter for this unit. However, we did want to use the particle model to explain chemical reactions and the conservation of matter.

Discussion of the design model.

We also used our design model (Figure 1) to systematically redesign the unit in order to encourage greater alignment. Because we could follow our alignment map (Appendix E) from the learning goal to the learning performances, learning tasks and assessments, it was explicit that if we changed any one component other pieces would need to be revised. For example, in concern 2 we discuss how our use of inquiry ability “explain” resulted in adding two standards, revising the learning performances, adding a learning task, altering the instructional sequence, and changing the assessments. Addressing this one concern resulted in revisiting 5 of the 6 steps in the learning goals and development stages of the design model. By revising these components in an iterative process and by using multiple sources of data, we strived to achieve greater alignment in the learning goals with the curriculum and assessment materials. To garner support for our assumptions, we had teachers react to the curriculum unit.

Second Enactment: Influence of the Revisions

The second version of our unit, expanded to 8 weeks, focused on all five of the selected chemistry standards about substances, properties, chemical reactions, the particle nature of matter, the conservation of mass on the macroscopic level, and the conservation of mass in terms of the particle nature of matter (Appendix A). During the 2002-2003 school year, we scaled the enactment of the unit to include more teachers, students, schools and sites than participated in the curricular unit the previous year (table 7).

Participants and Setting

Nine teachers enacted the unit in three different sites including the two locations from the first enactment (Urban A and Urban B) and one additional location (Large Town D). This enactment included 751 students in seven different schools.

Table 7: Teachers, students, and classrooms involved in the second enactment

Site	2002-2003 School Year			
	Urban A	Urban B	Large Town D	Total
Schools	4	2	1	7
Teachers	4	2	3	9
Classrooms	14	7	5	26
Students	471	209	71	751

Three of the four schools in Urban A were public middle schools, while the fourth school was a charter. Similar to the first enactment, the majority of the Urban A students come from lower to lower-middle income families (approximately half of the students lived in families that are at or below the poverty line), were largely minorities (over 90% are African Americans), and were mobile. The two schools in Urban B were public middle schools. The students in one school from Urban B came from lower to lower-middle income families with the majority of their families speaking Spanish as their primary language. The second school in Urban B consisted of an ethnically diverse population with students from lower-middle to middle income families. The three teachers in Large Town D taught at an independent middle school in a mid-size Midwest college town. The majority of these students were Caucasian and from middle to upper-middle income families.

Only students who completed both the pre- and posttest assessments were included in the analysis. The revised pre and posttest included more items than the first test and extended over two days before the unit and two days after the unit. Due to high absenteeism and mobility, especially in the urban classrooms, a number of students did not complete all four days of testing. Consequently, we did not analyze these students' pre and posttests.

Data Sources

Identical pre and posttest measures consisted of 30 multiple-choice and 6 open-ended items. Test items measured both the science content standards and scientific inquiry standards addressed in our learning performances. We scored and tallied multiple-choice responses for a maximum possible score of 30. We developed specific rubrics to score the six open-ended items with a total maximum score of 30 (See McNeill & Krajcik, in press). One rater initially scored the open-ended questions. We then randomly sampled 20% of the answer sheets and a second rater scored them. The average inter-rater reliability was above 85% for the six items.

Results

In order to evaluate whether our revision resulted in greater alignment with our learning goals and greater student learning, we analyzed the pre and posttest data from all three sites. Below, we discuss overall achievement, the chemical reaction content standard discussed in the revision process, scientific explanations and the problematic multiple choice item in concern #6.

Overall achievement by site. Table 8 contains the pretest and posttest data for the second enactment. Similar to the first enactment, students achieved significant learning gains, but the effect sizes in this enactment are for the most part larger. The effect sizes for the two urban sites are considerably larger than the effect sizes for those same two sites in the last enactment. The effect size for Large Town D was also much larger than the effect sizes in the previous enactment. Overall, the larger effect sizes suggest that this enactment resulted in greater student learning, possibly because of the revision to the curriculum materials and the more careful alignment of the learning goals with the materials and assessments.

Table 8: Enactment 2 test data by site

Site	Pretest <i>M</i> (<i>SD</i>) ^a	Posttest <i>M</i> (<i>SD</i>)	<i>t</i> -Value ^b	Effect Size ^c
Urban A (<i>n</i> =244)	15.93 (6.44)	30.83 (11.09)	26.65***	2.31
Urban B (<i>n</i> =162)	14.91 (6.78)	29.54 (10.18)	24.10***	2.16
Large Town D (<i>n</i> =71)	27.34 (7.13)	47.47 (6.72)	24.68***	2.82

^a Maximum score = 60

^b One-tailed paired *t*-test

^c Effect Size: Calculated by dividing the difference between posttest and pretest mean scores by the pretest standard deviation.

*** *p* < .001

Chemical reaction content standard achievement by site. To evaluate the effects of the revisions on the chemical reaction portion of the unit, we examined all the questions on the test that aligned with these learning performances. Table 9 shows the results of the chemical reaction standard for all of the items combined, the items that focused on macroscopic phenomena, and the items that focused on the particle nature of matter. Again, students achieved significant learning gains for this standard. The total effect size for the two urban sites was considerably

larger than the previous enactment. Large Town D also had a much larger effect size compared to the effect sizes for the two urban sites in the last enactment.

Since we added the particle nature of matter into the chemical reaction component of the unit, we were interested in whether the learning gains were different for the macroscopic phenomena compared to the particle nature of matter. We categorized all of the test items as either macroscopic or particle model (Table 9). While there are significant gains for both the macroscopic and the particle model, the effect sizes for the macroscopic phenomena are larger across all four sites. This suggests that we may want to revise both the instructional materials and assessment items that focus on the particle nature of matter.

Table 9: Enactment 2 data for chemical reactions by site

Site	Pretest M (SD) ^a	Posttest M (SD)	t -Value ^b	Effect Size ^c
Urban A ($n = 244$)				
Total Chem Rxn	6.84 (2.99)	11.52 (4.33)	18.78***	1.57
Macro	3.34 (2.14)	6.35 (3.04)	15.46***	1.40
Particle	3.50 (1.55)	5.18 (1.84)	14.74***	1.08
Urban B ($n = 162$)				
Total Chem Rxn	5.96 (2.91)	11.36 (4.03)	18.43***	1.86
Macro	2.78 (2.15)	6.61 (2.94)	15.89***	1.78
Particle	3.18 (1.52)	4.75 (1.72)	10.09***	1.03
Large Town D ($n = 71$)				
Total Chem Rxn	9.92 (3.28)	18.04 (2.63)	20.15***	2.48
Macro	5.02 (2.74)	11.39 (2.02)	18.75***	2.32
Particle	4.90 (1.51)	6.65 (1.22)	9.03***	1.16

^a Maximum score: Total = 26, Macro = 17.5, Particle = 8.5

^b One-tailed paired t -test

^c Effect Size: Calculated by dividing the difference between posttest and pretest mean scores by the pretest standard deviation.

- $p < .05$; ** $p < .01$; *** $p < .001$

Overall, we observed larger learning gains for the chemical reaction items in the second enactment compared to the first enactment. This suggests that the revision of the materials resulted in greater alignment and support for these learning goals.

Revisiting concern #2: students' construction of scientific explanations. We analyzed if the changes in the unit resulted in greater student understanding of scientific explanations, particularly the reasoning component. Table 10 shows the results from this analysis.

Table 10: Enactment 2 data for scientific explanations

Site	Pretest M (SD) ^a	Posttest M (SD)	t -Value ^b	Effect Size ^c
Urban A ($n = 244$)				
Total	1.25 (1.64)	3.13 (2.55)	11.41***	1.15
Claim	0.73 (1.00)	1.42 (1.25)	7.68***	0.69
Evidence	0.42 (0.733)	1.00 (0.98)	8.77***	0.79
Reasoning	0.10 (0.29)	0.71 (0.97)	10.02***	2.10
Urban B ($n = 162$)				
Total	0.71 (1.39)	3.13 (2.16)	13.84***	1.74
Claim	0.43 (0.86)	1.66 (1.17)	11.23***	1.43
Evidence	0.23 (0.52)	0.67 (0.80)	6.73***	0.85
Reasoning	0.05 (0.27)	0.80 (0.97)	10.19***	2.78
Large Town D ($n = 71$)				
Total	3.23 (2.52)	6.89 (2.26)	11.42***	1.45
Claim	1.68 (1.28)	2.89 (0.89)	8.10***	0.95
Evidence	1.15 (1.15)	2.08 (1.11)	5.45***	0.81
Reasoning	0.40 (0.71)	1.92 (0.95)	11.68***	2.14

^a Maximum score: Total = 10, Claim = 3.3, Evidence = 3.3, Reasoning = 3.3

^b One-tailed paired t -test

^c Effect Size: Calculated by dividing the difference between posttest and pretest mean scores by the pretest standard deviation.

- $p < .05$; ** $p < .01$; *** $p < .001$

Overall, students achieved significant gains for all three components of scientific explanation. Again, the effect sizes in the second enactment were larger than the first (see table 5). Furthermore, the significance of the claim and reasoning learning gains increased compared to the results during the last enactment. Across the three sites, the total effect size, claim and reasoning were larger than the effect size the previous year. During the revision of the unit, we particularly targeted the reasoning component of explanation because students had difficulty with this both on the pre and posttest and the student artifacts. The results from the second enactment suggest that students had a greater understanding of this component.

Revisiting concern #6: students thought mixtures were a chemical reaction. On the revised pre and posttest a similar question was included to the one students had difficulty with in the first enactment about mixtures. The question asked, “Which will produce new substances?” The following choices were provided: a. hammering a piece of metal, b. burning a candle, c. heating water until it evaporates, and d. dissolving lemonade powder in a liquid. The frequency of student choices for the pretest was very similar in both enactment #1 and enactment #2 (see table 11). In the first enactment, more students selected dissolving lemonade after the instructional unit than before. In the second enactment, the number of students selecting dissolving lemonade did decrease. While the choice decreased, 39% of students on the posttest are still selecting dissolving lemonade. Helping middle school students distinguish between a dissolving and chemical reaction is very challenging.

Table 11: Frequencies for Student Choices on Multiple Choice Item by enactment

	Enactment #1 (n=89)		Enactment #2 (n=474)	
	pretest	posttest	pretest	posttest
Stretching a rubber band/Hammering metal	2.2%	1.1%	6.96%	1.9%
Burning a Candle	15.7%	16.9%	22.2%	39.0%
Heating Water	27.0%	14.6%	23.6%	19.6%
Dissolving Lemonade	55.1%	67.4%	47.4%	39.0%

Adding the lesson on mixtures, appears to have helped students with this concept, yet it continues to be a difficult area. In our next round of revision, we plan to revisit this section of the instructional materials in order to further address this area of concern.

Conclusion

The analyses of the second enactment reveal greater student learning gains on the pre and posttest for the total score, as well as specifically for the chemical reaction and scientific explanation learning goals. Greater alignment of learning goals with instructional materials and assessment measures, which our design model encouraged, provides one possible explanation to account for student gains in the pre- to posttest. Our work presents three important lessons that provide insight into the importance of how to align learning goals with teaching and student materials and assessments, and the importance of clearly specifying learning goals in order to promote learning of content and inquiry goals. We discuss these lessons below.

First, the iterative design model (Figure 1) can help explain the improvement in student learning seen during the second enactment. Like Linn and her colleagues, we used an iterative approach to curriculum development to promote student learning. Linn's Computers as Learning Partners (CLP) project serves as an excellent model of how the iterative revision can result in improving student understanding of challenging science content. In the case of CLP, Linn and colleagues demonstrated improved learning of a challenging content – heat energy and temperature (Linn & Hsi, 2000). Our work takes Linn's important contribution a step further because we show how you can match learning goals to standards. A major aspect of our design involved first unpacking standards and then translating these standards into learning performances.

Results from the second enactment suggest that greater student learning of both science content standards and inquiry standards resulted from the revision of student and teacher materials. This iterative design model (See Figure 1) is critical in aligning learning goals with teacher and student materials, and assessments. Our model promotes this alignment by making these links to the learning goals explicit as well as by encouraging multiple iterations of the design process. By tracing each change to the other parts of the materials, we created consistency across the unit including the assessments, which are often a neglected portion of the design process. This alignment process is consistent with Wilson and Bertenthal (National Research Council, 2006) who argue for the importance of aligning learning goals with instruction and assessment. Yet we provide a concrete example rather than just a recommendation. Moreover, we provide a model for how other curriculum designers including teachers can operationalize this recommendation

Second, the revision process revealed the importance of using multiple data sources with the use of analytical and empirical analyses, during the development of the instructional materials to further ensure alignment of the learning goals with teacher and student materials and assessments. Each data source provided a unique perspective as well as reinforced the importance of concerns identified from other data sources. There are trade-offs in design choices in both the initial design and revision of the materials. We felt these different perspectives allowed us to make more informed decisions regarding the pros and cons behind each decision. Although such a process is challenging and requires a significant time commitment, it provides a model for developing materials that can help all students develop deep understanding of important and challenging science content because of the tight alignment that results between learning goals, instructional materials and assessments.

Third, transforming the science content standards into learning performances based on research from learning and instruction served as an important step in this design process. Many researchers in the field have argued for the importance of clearly specifying what we want students to know (Perkins, Crismond, Simmons & Unger, 1995; Wiggins & McTighe, 1998). We translate this idea into practice by creating learning performances. Creating learning performances forced us to explicitly state what it means for a student to “know” science content. We believe that it is important to address the content and inquiry standards simultaneously because the knowing of science cannot be separated from the doing of science (Marx et al., 1997). Writing and revising the learning performances allowed us to form a clearer image of what we expected of students, and what the learning tasks and assessment measures needed to include. In essence, creating learning performance encouraged greater alignment of our learning goals (e.g. learning performances), learning tasks and assessment measures. Furthermore, the learning performances allowed us to look at the same content across different inquiry practices and the same inquiry practices across different content in order to create a more complete picture of a student’s understanding. Our work on learning performances provides a model of how other designers can specify what students should know.

The work we report here supports the use of a learning-goal driven model to develop instructional materials that align with standards and assessments. Such alignment should result in increased students learning.

References

- American Association for the Advancement of Science. (1990). *Science for All Americans: A project 2061 report on the literacy goals in science, mathematics, and technology*. New York: Oxford University Press.
- American Association for the Advancement of Science. (1993). *Benchmarks for Science Literacy*. New York: Oxford University Press.
- American Association for the Advancement of Science. (2001). *Atlas of Science Literacy*. Washington, DC: American Association for the Advancement of Science & National Science Teachers Association.
- Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. New York: Longman.
- Blumenfeld, P., Fishman, B., Krajcik, J., Marx, R. W., & Soloway, E. (2000). Creating Useable Innovations in Systemic Reform: Scaling-up technology-embedded project-based science in urban schools. *Educational Psychologist*, 35(3), 149-164.
- Bransford, J., Brown, A., & Cocking, R. (Eds.). (2000). *How People Learn: Brain, Mind, Experience and School*. Washington D.C.: National Academy Press.
- Donovan, M. S. & Bransford, J. D. (Eds.), (2005). *How students learn: Science in the classroom*. Washington, DC: National Academy Press.
- Driver, R., Asoko, H., Leach, J., Mortimer, E., & Scott, P. (1994). Constructing scientific knowledge in the classroom. *Educational Researcher*, 23 (7), 5-12.
- Driver, R., Squires, A., Rushworth, P., and Wood-Robinson, V. (1994) *Making Sense of Secondary Science: Research into Children's Ideas*. London: Routledge Falmer.
- Glasser & Chi, 1988
- Duschl, R. A., & Hamilton, R. J. (1998). Conceptual Change in Science and in the Learning Science of Science. In B. F. K. Tobin (Ed.), *International Handbook of Science Education* (pp. 1047-1065). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Harris, C. J., McNeill, K. L., Lizotte, D. J, Marx, R. W., & Krajcik, J. (in press). Aligning standards-based assessment and curriculum materials for teaching inquiry science. *PEERs Matter*.
- Kesidou, S., & Roseman J. E. (2002). How well do Middle School Science Programs Measure Up? Findings from Project 2061's Curriculum Review. *Journal of Research in Science Teaching*. 39(6). 522-549.
- Krajcik, J., Blumenfeld, P., Marx, R., & Soloway, E. (2000). Instructional, Curricular, and Technological Supports for Inquiry in Science Classrooms. In J. Minstrell & E. v. Zee (Eds.), *Inquiring into Inquiry Learning and Teaching in Science* (pp. 283-315). Washington D.C.: AAAS.
- Kuhn, T. S. (1970). *The Structure of Scientific Revolutions*. Chicago, IL: University of Chicago Press.
- Lee, O., Eichinger, D. C., Anderson, C. W., Berkheimer, G. D., & Blakeslee, T. D. (1993). Changing middle school students' conceptions of matter and molecules. *Journal of Research in Science Teaching*, 30 (3), 249-270.
- Marx, R. W., Blumenfeld, P. C., Krajcik, J. S. & Soloway, E. (1997). Enacting Project-Based Science. *The Elementary School Journal*. 97(4). 341-358.
- McNeill, K. L. & Krajcik, J. (in press). Middle school students' use of appropriate and inappropriate evidence in writing scientific explanations. In Lovett, M & Shah, P (Eds.)

- Thinking with Data: the Proceedings of the 33rd Carnegie Symposium on Cognition. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- McNeill, K. L., Lizotte, D. J., Krajcik, J., & Marx, R. W. (in press). Supporting students' construction of scientific explanations by fading scaffolds in instructional materials. *The Journal of the Learning Sciences*.
- Moje, E.B., Peek-Brown, D., Sutherland, L.M., Marx, R.W., Blumenfeld, P., Krajcik, J. (2004). Explaining explanations: Developing scientific literacy in middle-school project-based science reforms. In D. Strickland & D. E. Alvermann, (Eds.), *Bridging the Gap: Improving Literacy Learning for Preadolescent and Adolescent Learners in Grades 4-12*. NY: Teachers College Press.
- National Research Council. 1996. *National Science Education Standards* Washington, DC: National Academy Press.
- National Research Council (2006). *Systems for state science assessment. Committee on Test Design for K-12 Science Achievement*. M. R. Wilson and M. W. Bertenthal, eds. Board on Testing and Assessment, Center for Education, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.
- Pellegrino, J. W., Chudowsky, N., & Glaser, R. (Eds.). (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academy Press.
- Perkins, D., D. Crismond, et al. (1995). Inside understanding. *Software goes to school: Teaching for understanding with new technologies*. D. Perkins, J. Schwartz, M. West and M. Wiske. New York, Oxford University Press: 70-87.
- Ruiz-Primo, M. A., Shavelson, R. J., Hamilton, L., & Klein, S. (2002). On the evaluation of systemic science education reform: Searching for instructional sensitivity. *Journal of Research in Science Teaching*, 39 (5), 369–393.
- Singer, J., Marx, R., Krajcik, J., & Chambers, J. (2000). Constructing Extended Inquiry Projects: Curriculum Materials for Science Education. *Educational Psychologist*. 35(3), 165-178.
- Smith, C. L., Wiser, M., Anderson, C. W., Krajcik, J., (in press). Implications of Research on Children's Learning for Standards and Assessment: A Proposed Learning Progression for Matter and the Atomic Molecular. *Theory Measurement: Interdisciplinary Research and Perspectives*.
- Wiggins, G., & McTighe, J. (1998). *Understanding by Design*. Alexandria, VA: Association for Supervision and Curriculum Development.

Figure 1: Learning-Goals-Driven Design Model

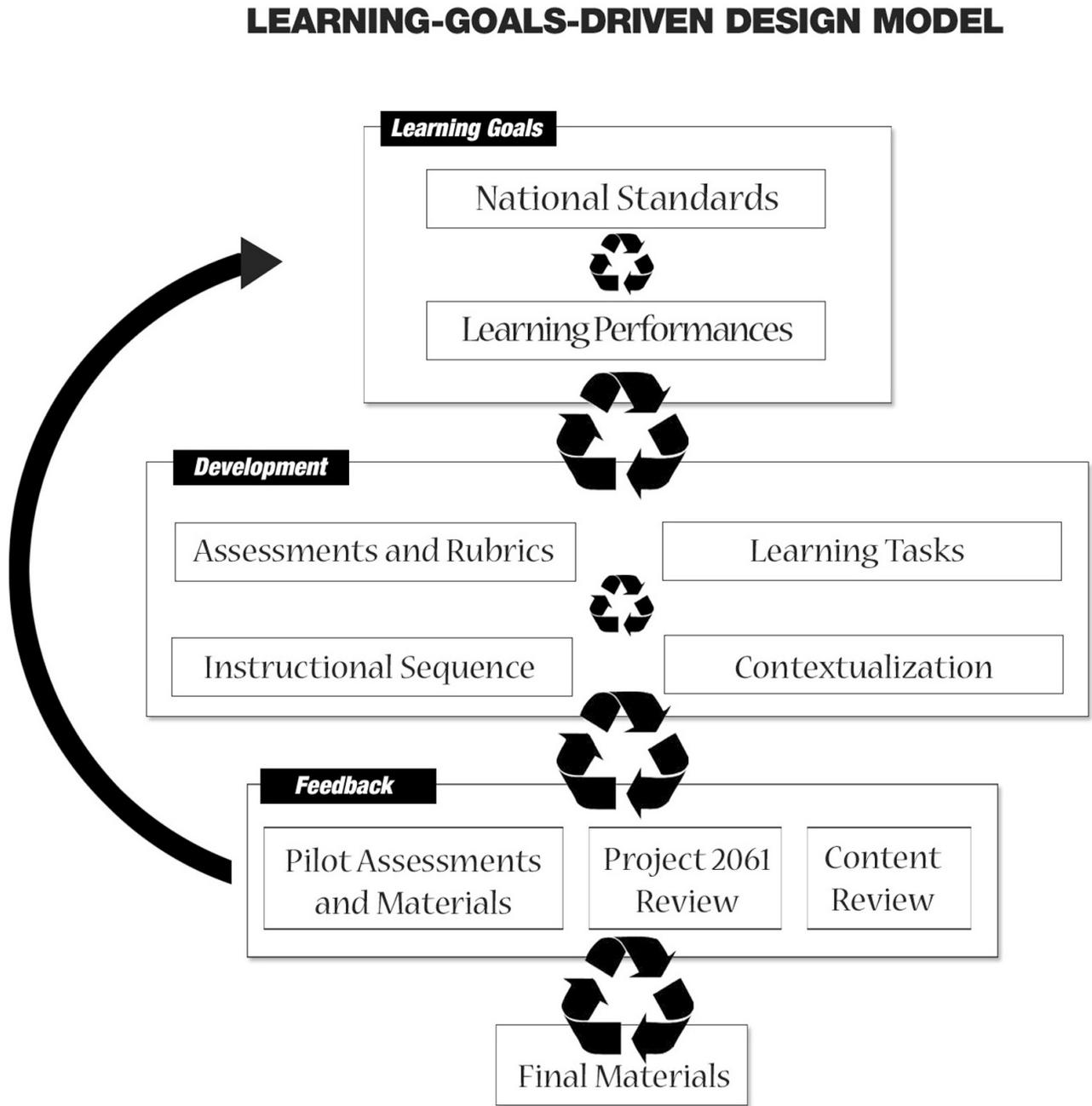
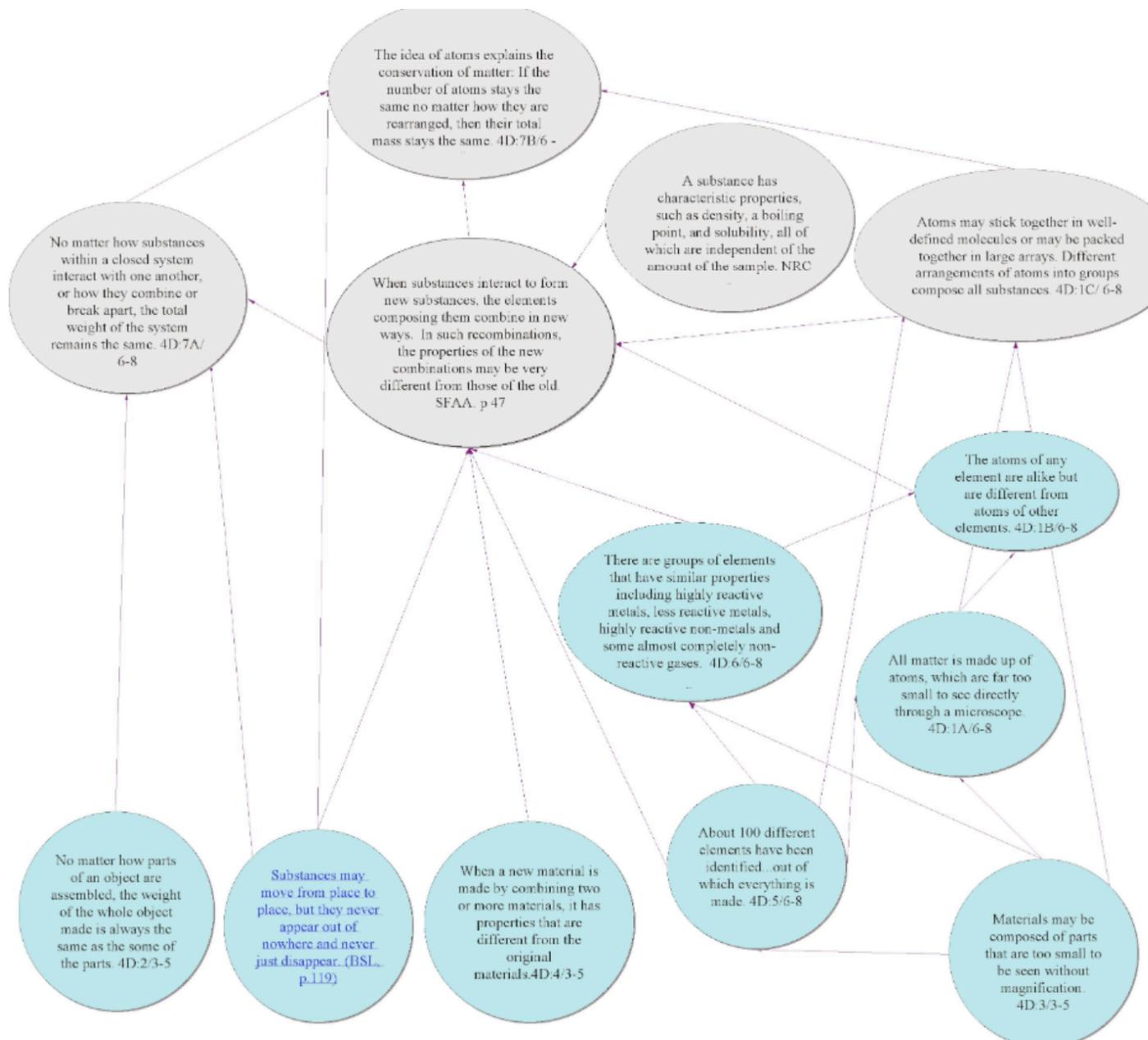


Figure 2: Concept Map of Key Standards



Appendix A: Unpacking the National Standards

Standard	“Unpacking” the Standard
<p>SFAA: When substances interact to form new substances, the elements composing them combine in new ways. In such recombinations, the properties of the new combinations may be very different from those of the old (AAAS, 1990, p.47).</p>	<p>Substances have distinct properties and are made of one material throughout. A chemical reaction is a process where new substances are made from old substances. One type of chemical reaction is when two substances are mixed together and they interact to form new substance(s). The properties of the new substance(s) are different from the old substance(s). When scientists talk about “old” substances that interact in the chemical reaction, they call them reactants. When scientists talk about new substances that are produced by the chemical reaction, they call them products.</p>
<p>4D7 – Part I No matter how substances within a closed system interact with one another, or how they combine or break apart, the total weight of the system remains the same (AAAS, 1993).</p>	<p>A closed system is when matter cannot enter or leave a physical boundary. Regardless of how materials interact with each other or change by breaking apart and forming new combinations in a closed system, the total mass of all the material in the system remains the same. The amount of material in our system is represented by the mass of the system. In this case, we are interpreting weight as mass. A common misconception of students is to use mass and weight to have the same meaning. We believe that we need to be consistent in 4D7–Part I and 4D7-Part II. Therefore we are using the term mass in both Part I and Part II.</p>
<p>4D: 1 – Part II Atoms may stick together in well-defined molecules or may be packed together in large arrays. Different arrangements of atoms into groups compose all substances (AAAS, 1993).</p>	<p>Atoms can be arranged in particular ways including the formation of discrete molecules and arrays. A molecule is made up of atoms stuck together in a certain arrangement. An array has repeated patterns of atoms. The different arrangements of atoms give materials different properties. Materials with unique properties are different substances.</p>
<p>4D7 - Part II The idea of atoms explains the conservation of matter: If the number of atoms stays the same no matter how they are rearranged, then their total mass stays the same (AAAS, 1993).</p>	<p>The conservation of matter states that regardless of how substances interact with each other in a closed system, the total mass of all the substances in the system remains the same (4D 7-Part I). The majority of substances are made of molecules that are composed of atoms. The reason that the conservation of matter occurs is because the number of atoms of each element in the system stays the same. Regardless of how atoms interact (by breaking apart and reforming new molecules or new arrays) with each other in a closed system, their total mass in the system remains the same.</p>
<p>ADDED DURING DESIGN PROCESS: B 5-8: 1A A substance has characteristic properties, such as density, a boiling point, and solubility, all of which are independent of the amount of the sample (NRC, 1996, p.154)</p>	<p>Substances have distinct properties that can be used to distinguish and separate one substance from another. Properties such as density, melting point, and solubility, describe the unique characteristics of substances. Density is the mass contained within a unit volume. Melting point is the temperature at which a solid changes to a liquid. Solubility is the ability of a solid to dissolve in a liquid.</p>

Appendix B: Enactment #1 - Learning Performances

Standard	Learning Performance
<p>B 5-8: 1A A substance has characteristic properties, such as density, a boiling point, and solubility, all of which are independent of the amount of the sample (NRC, 1996, p.154)</p>	LP1 – Students identify and describe observable properties of substances.
	LP2- Students measure the density, melting point, and solubility of substances.
	LP3 – Students explain that properties are unique characteristics that help identify. These properties do not change regardless of the amount of the substance.
	LP4 – Students explain that substances have distinct properties that can be used to distinguish one substance from another.
<p>SFAA: When substances interact to form new substances, the elements composing them combine in new ways. In such recombinations, the properties of the new combinations may be very different from those of the old (AAAS, 1990, p.47).</p>	LP5 - Students identify and describe the properties of substances before and after a chemical reaction.
	LP6 - When various substances come in contact with each other, students identify whether a chemical reaction has occurred, provide evidence, and explain.
	LP7 - Students design an experiment to determine whether a chemical reaction occurred. They make predictions about what will happen, carry out their investigation, and explain how the evidence supports their conclusion that a chemical reaction either did or did not occur.
	LP8– Students explain that a chemical reaction is a process where a new substance is made from an old substance.

Appendix C: Enactment #1 – Project Calendar

PROJECT CALENDAR: HOW CAN I MAKE NEW STUFF FROM OLD STUFF?	
Learning Set One: How is stuff the same and different?	
2 Class Periods	Lesson #1 – How is this stuff the same or different? <ul style="list-style-type: none">Describe objects from the classroomDescribe unknown stuff (fat and soap)Describe Box of stuff used to introduce the concepts of “substance” and “property”
2 Class Periods	Lesson #2 – What are the properties of this stuff? <ul style="list-style-type: none">Demonstrate density using two different metal blocksCalculate density of two different sized wooden blocksCalculate density of fat and soap
2 Class Periods	Lesson #3 – What else is different about this stuff? <ul style="list-style-type: none">Demonstrate solubility and melting point using butter and margarine.Determine solubility of fat and soapDetermine melting point of fat and soap
Learning Set Two: What happens when you combine stuff?	
2 Class Periods	Lesson #4 – What happens to properties when I combine stuff? <ul style="list-style-type: none">Complete chemical reaction in the baggie (phenol red, road salt, baking soda & sugar) used to introduce the concept of “chemical reaction”Redesign the investigation to determine what combination of ingredients caused a specific change or indicator.
2 Class Periods	Lesson #5 – Is this new stuff? <ul style="list-style-type: none">Investigate whether boiling is a chemical reactionInvestigate whether the electrolysis of water is a chemical reaction
2 Class Periods	Lesson #6 – How can I make new stuff from this stuff? <ul style="list-style-type: none"><i>Combine glue, water, and sodium borate solution to create a new substance (slime).</i><i>Create a new recipe for the “bounciest” slime and design a procedure to test the bounce.</i>
3 Class Periods	Lesson #7 – How can I make fat from soap? <ul style="list-style-type: none">Conduct experiment to make soap from fat. Discuss student reader about the history of soap making.
2 Class Periods	Lesson #8 – How can I change the driving question? <ul style="list-style-type: none">Decide whether or not to change the wording of the driving question to make it more scientific.Present new driving question and reasons for any changes to the class.

Appendix D: Base Explanation Rubric

<u>Component</u>	<u>Level</u>		
	<u>0</u>	<u>1</u>	<u>2</u>
<u>Claim</u> – An assertion or conclusion that answers the original question.	<u>Does not make a claim, or makes an inaccurate claim.</u>	<u>Makes an accurate but incomplete claim.</u>	<u>Makes an accurate and complete claim.</u>
<u>Evidence</u> – Scientific data that supports the claim. The data needs to be appropriate and sufficient to support the claim.	<u>Does not provide evidence, or only provides inappropriate evidence (Evidence that does not support claim).</u>	<u>Provides appropriate, but insufficient evidence to support claim. May include some inappropriate evidence.</u>	<u>Provides appropriate and sufficient evidence to support claim.</u>
<u>Reasoning</u> – A justification that links the claim and evidence and shows why the data counts as evidence to support the claim by using the appropriate and sufficient scientific principles.	<u>Does not provide reasoning, or only provides reasoning that does not link evidence to claim.</u>	<u>Provides reasoning that links the claim and evidence. Repeats the evidence and/or includes some scientific principles, but not sufficient.</u>	<u>Provides reasoning that links evidence to claim. Includes appropriate and sufficient scientific principles.</u>

Appendix D: Specific Rubrics

Specific Rubric for Substance and Property Scientific Explanation

<u>Component</u>	<u>Level</u>		
<p><u>Claim –</u> A statement or conclusion that answers the original question/problem</p>	<u>0</u>	<u>1</u>	<u>2</u>
	<p>Does not make a claim, or makes an inaccurate claim.</p> <p>-----</p> <p>States none of the liquids are the same or specifies the wrong solids.</p>	<p>Makes an accurate but incomplete claim.</p> <p>-----</p> <p>Vague statement, like “some of the liquids are the same.”</p>	<p>Makes an accurate and complete claim.</p> <p>-----</p> <p>Explicitly states “Liquids 1 and 4 are the same substance.”</p>
<p><u>Evidence –</u> Scientific data that supports the claim. The data needs to be appropriate and sufficient to support the claim.</p>	<u>0</u>	<u>1 & 2</u>	<u>3</u>
	<p>Does not provide evidence, or only provides inappropriate evidence (Evidence that does not support claim).</p> <p>-----</p> <p>Provides inappropriate data, like “the mass is the same” or provides vague evidence, like “the data table is my evidence.”</p>	<p>Provides appropriate, but insufficient evidence to support claim. May include some inappropriate evidence.</p> <p>-----</p> <p>Provides 1 or 2 of the following pieces of evidence: the density, melting point, and colors of liquids 1 and 4 are the same. May also include inappropriate evidence, like mass.</p>	<p>Provides appropriate and sufficient evidence to support claim.</p> <p>-----</p> <p>Provides all 3 of the following pieces of evidence: the density, melting point, and colors of liquids 1 and 4 are the same.</p>
<p><u>Reasoning –</u> A justification that links the claim and evidence and includes appropriate and sufficient scientific principles to defend the claim and evidence.</p>	<u>0</u>	<u>1, 2 & 3</u>	<u>4</u>
	<p>Does not provide reasoning, or only provides reasoning that does not link evidence to claim.</p> <p>-----</p> <p>Provides an inappropriate reasoning statement like “they are like the fat and soap we used in class” or does not provide any reasoning.</p>	<p>Repeats evidence and links it to the claim. May include some scientific principles, but not sufficient.</p> <p>-----</p> <p>Repeats the density, melting point, and colors are the same and states that this shows they are the same substance. Or provides an incomplete generalization about properties, like “mass is not a property so it does not count.”</p>	<p>Provides accurate and complete reasoning that links evidence to claim. Includes appropriate and sufficient scientific principles.</p> <p>-----</p> <p>Includes a complete generalization that density, melting point, and color are all properties. Different substances have different properties. Since liquids 1 and 4 have different properties there are different substances.</p>

Appendix D: Specific Rubrics

Specific Rubric for Chemical Reaction Scientific Explanation

<u>Component</u>	<u>Level</u>		
<u>Claim –</u> A statement or conclusion that answers the original question/problem	<u>0</u>	<u>Does not apply to this learning task.</u>	<u>1</u>
	<u>Does not make a claim, or makes an inaccurate claim.</u> ----- <u>States that a chemical reaction did not occur.</u>		<u>Makes an accurate and complete claim.</u> ----- <u>States that a chemical reaction did occur.</u>
<u>Evidence –</u> Scientific data that supports the claim. The data needs to be appropriate and sufficient to support the claim.	<u>0</u>	<u>1 & 2</u>	<u>3</u>
	<u>Does not provide evidence, or only provides inappropriate evidence (Evidence that does not support claim).</u> ----- <u>Provides inappropriate data, like “the mass and volume changed” or provides vague evidence, like “the data shows me it is true.”</u>	<u>Provides appropriate, but insufficient evidence to support claim. May include some inappropriate evidence.</u> ----- <u>Provides 1 or 2 of the following pieces of evidence: Butanic acid and butanol have different solubilities, melting points, and densities compared to Layer A and Layer B. May also include inappropriate evidence, like mass or volume.</u>	<u>Provides appropriate and sufficient evidence to support claim.</u> ----- <u>Provides all 3 of the following pieces of evidence: Butanic acid and butanol have different solubilities, melting points, and densities compared to Layer A and Layer B. May also include inappropriate evidence, like mass.</u>
<u>Reasoning –</u> A justification that links the claim and evidence and includes appropriate and sufficient scientific principles to defend the claim and evidence.	<u>0</u>	<u>1, 2, 3 & 4</u>	<u>5</u>
	<u>Does not provide reasoning, or only provides reasoning that does not link evidence to claim.</u> ----- <u>Provides an inappropriate reasoning statement like “a chemical reaction did not occur because Layers A and B are not substances” or does not provide any reasoning.</u>	<u>Repeats evidence and links it to the claim. May include some scientific principles, but not sufficient.</u> ----- <u>Repeats the solubility, melting point, and density changed, which show a reaction occurred. Or provides either A or B:</u> A. <u>A chemical reaction creates new or different substances</u> OR B. <u>Different substances have different properties.</u>	<u>Provides accurate and complete reasoning that links evidence to claim. Includes appropriate and sufficient scientific principles.</u> ----- <u>Includes a complete generalization that:</u> A. <u>A chemical reaction creates new or different substances</u> AND B. <u>Different substances have different properties.</u>

Appendix E: Enactment 1 - Curriculum and assessment alignment map

Lesson Title	Brief Description	Assessment Items^a	Learning Performances^b	Standard^c
Lesson 1: How is this stuff the same or different?	Students make careful observations, write descriptions, and compare similarities and differences of two unknowns (lard and soap).	MC: 1, 5, 17	LP1	B 5-8: 1A A substance has characteristic properties, such as density, a boiling point, and solubility, all of which are independent of the amount of the sample (NRC, 1996, p.154)
Lesson 2: What are the properties of this stuff?	Students measure and compare densities of lard and soap.	MC: 2 MC: 14	LP2 LP3 LP4	
Lesson 3: What else is different about this stuff?	Students measure and compare melting points and solubility of lard and soap.	MC: 7, 16, 18 OE: 23		
Lesson 4: What happens to properties when I combine stuff?	Students create a chemical reaction by combining calcium chloride, baking soda, sugar, and phenol red solution. They design an experiment to determine the cause of one change in property or indicator.	MC: 11, 15, 20 MC: 13 OE: 22 OE: 24 MC: 3, 8, 12	LP5	SFAA: When substances interact to form new substances, the elements composing them combine in new ways. In such recombinations, the properties of the new combinations may be very different from those of the old (AAAS, 1990, p.47).
Lesson 5: Is this new stuff?	Students measure and compare properties of water that has been boiled versus water that has been subjected to electrolysis.		LP6	
Lesson 6: How can I make new stuff from this stuff?	Students examine the properties of glue, water, and sodium borate solution separately and then after they react to make slime.		LP7	
Lesson 7: How can I make soap from fat?	Students make soap from lard and sodium hydroxide solution. They measure and compare densities, melting points, and solubilities of their own soap with commercial soap.		LP8	

^aAssessment Items: MC = multiple choice, OE = open ended.

^bSee Appendix B for description of learning performances.

^cSee Appendix A for an unpacking of the knowledge specified in the standards.

Appendix F: Revised Learning Performances:

Content Standard	Inquiry Standard	Learning Performance
<p>SFAA: When substances interact to form new substances, the elements composing them combine in new ways. In such recombinations, the properties of the new combinations may be very different from those of the old (AAAS, 1990, p.47).</p>	Develop descriptions...using evidence. (NRC, 1996, A: 1/4, 5-8)	LP 6 - Students identify and describe the properties of substances before and after a chemical reaction.
	Develop...explanations... using evidence. (NRC, 1996, A: 1/4, 5-8)	LP7 - Students create scientific explanations stating a claim whether a chemical reaction occurred, evidence in the form of properties, and reasoning that a chemical reaction is a process where old substances interact to form new substances with different properties from the old substances.
	Think critically and logically to make the relationships between evidence and explanation. (NRC, 1996, A: 1/5, 5-8)	
	Design and conduct a scientific investigation. (NRC, 1996, A: 1/2, 5-8)	LP8 - Students design an experiment to determine what combination of a given number of substances causes a chemical reaction. They make predictions about what will happen, carry out their investigation, and collect evidence. They determine whether the evidence supports their prediction that a chemical reaction either did or did not occur.
	When similar investigations give different results, the scientific challenge is to judge whether the differences are trivial or significant... (AAAS, 1993, 1A: 1, 6-8)	LP9 - Students compare and contrast two or more processes for the same substance to determine whether the processes create the same products. They determine if any of the processes is a chemical reaction.
	Models are often used to think about processes that happen... too quickly, or on too small a scale to observe directly... (AAAS, 1993, 11B: 1, 6-8)	LP10 - Students use particle models to represent what happens to atoms and molecules during a chemical reaction demonstrating that the atoms recombine and stick together in different arrangements to form atoms or new molecules.
Develop...models using evidence. (NRC, 1996, A: 1/4, 5-8)		
Models are often used to think about processes that happen... too quickly, or on too small a scale to observe directly (AAAS, 1993, 11B: 1, 6-8)	LP11 - Given a representation of the atoms and molecules in a chemical reaction, mixture, or phase change, students describe whether or not a chemical reaction occurred based on whether the substances interact and their atoms recombine and stick together in different arrangements to form atoms or new molecules.	
Develop...models using evidence. (NRC, 1996, A: 1/4, 5-8)		