Supporting Students' Construction of Scientific Explanations

By Fading Scaffolds in Instructional Materials

Katherine L. McNeill, David J. Lizotte, & Joseph Krajcik

University of Michigan

Ronald W. Marx

University of Arizona

contact info:

Center for Highly Interactive Classrooms, Curricula & Computing in Education

610 E. University Ave., Ann Arbor, MI, 48109-1259

734-647-4226

kmcneill@umich.edu

Reference as:

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.*

Abstract

The purpose of this study is to determine whether providing students with continuous written instructional support or fading written instructional support (scaffolds) better prepares students to construct scientific explanations when they are no longer provided with support. We investigated the influence of scaffolding on 331 7th grade students' writing of scientific explanations during an 8-week project-based chemistry unit in which the construction of scientific explanations is a key learning goal. The unit makes an instructional model for explanation explicit to students through a focal lesson and reinforces that model through subsequent written support for each investigation. Students received one of two treatments in terms of the type of written support: Continuous, involving detailed support for every investigation, or Faded, involving less support over time. Our analyses showed significant learning gains for students for all components of scientific explanation (i.e. claim, evidence, and reasoning). Yet on posttest items lacking scaffolds, the Faded group gave stronger explanations in terms of their reasoning compared to the Continuous group. Fading written scaffolds better equipped students to write explanations when they were not provided with support.

Supporting Students' Construction of Scientific Explanations By Fading Scaffolds in Instructional Materials

Ultimately the goal of classroom science is to help all students become scientifically literate to encourage greater public understanding in a science infused world (American Association for the Advancement of Science, 1993; National Research Council, 1996). This type of literacy requires that students be able to participate in science discourses and practices (McGinn & Roth, 1999). "Learning science involves young people entering into a different way of thinking about and explaining the natural world; becoming socialized to a greater or lesser extent into the practices of the scientific community with its particular purposes, ways of seeing, and ways of supporting its knowledge claims" (Driver, Asoko, Leach, Mortimer, & Scott, 1994, p. 8). Engaging in these science discourses and practices is often difficult for middle-school students because they are new to the students (Krajcik, Blumenfeld, Marx, Bass, & Fredericks, 1998). Without support for learning these new ways of knowing, doing, and talking science, students may not relate to science and even actively resist learning it (Lee & Fradd, 1998).

A core practice of science is the construction of arguments or explanations including weighing evidence, interpreting text, and evaluating claims (Driver, Newton, & Osborne, 2000). Scientific explanations frame the goal of inquiry as understanding natural phenomena, and articulating and convincing others of that understanding (Sandoval & Reiser, 2004), in contrast to the view of classroom science as simply learning facts about the natural world. Having students engage in this type of explanation construction may change or refine their image of science (Bell & Linn, 2000) as well as enhance students' understandings of the science content (Zohar & Nemet, 2002). Although scientific explanations are important for classroom science, they are frequently omitted from classroom practice (Kuhn, 1993; Newton, Driver & Osborne 1999). Furthermore, when students do engage in explanation or argumentation they often have difficulty articulating and justifying their claims (Sadler, 2004). Because of these challenges, students need to be explicitly taught about scientific explanation in order to be successful in this practice (Osborne, et al, 2004). We are interested in how to best assist students in their construction of scientific explanations. Our work focuses on an 8-week project-based chemistry curriculum designed to support 7th grade students in their understanding of scientific explanations. Our study investigates whether fading written instructional supports (scaffolds) or providing students with continuous written instructional support better equips students to construct explanations when they are not provided with support.

Conceptual Framework

In order to frame our research study, we first discuss recent issues in the explanation and argumentation literature as well as the scaffolding literature. We begin by describing common student difficulties with explanation and argumentation. Then we describe the instructional model we developed to help both middle school students and teachers with this important and difficult inquiry practice. Specifically, we are interested in the implementation of this instructional model and the most effective way to incorporate the model into the design of science curriculum materials. We discuss the importance of scaffolding student learning and current questions in educational literature in terms of the fading of written supports. We use this literature to inform the design of the instructional materials and written supports as well as to frame our research question of whether providing continuous written instructional support or fading written instructional support (scaffolds) better prepare students to construct explanations when they are no longer provided with support.

Student Difficulties Constructing Scientific Explanations

Our work draws on both the explanation and argumentation literature in science education. Explanations often refer to how or why something happens (Chin & Brown, 2000). Instead of simply identifying that a phenomena occurred, scientists try to explain phenomena by determining how or why they occur and the conditions and consequences of the observed event (Nagel, 1961). An argument is an assertion with a justification (Kuhn, 1991). Argumentation is typically viewed as a verbal (written or oral) and social activity aimed at justifying or defending a standpoint for an audience (van Eemeren, et al., 1996). An argument can be constructed to explain a scientific phenomena. For example, a scientist could construct an argument to explain why biodiversity is decreasing in a particular ecosystem. But an argument can also be constructed for other reasons, such as to support an individual's opinion or belief. Our goal is to help students construct scientific explanations about phenomena where they justify their claims using appropriate evidence and scientific principles. Previous research examining students' explanations and arguments suggests that students have difficulty articulating and defending their claims (Sadler, 2004).

Evidence. In explaining scientific phenomena, students need to be able to gather, select and use data as evidence to support their claims. Yet students often have difficulty with this complex task. Students struggle to understand what counts as evidence (Sadler, 2004) and to use appropriate evidence (Sandoval, 2003). Not only students, but adults with a range of educational experiences and expertise also have difficulty distinguishing evidence from theory and using evidence to support their claims (Kuhn, 1991). Instead of using evidence, students often rely on their personal views to draw conclusions (Hogan & Maglienti, 2001). Students may draw on other knowledge and beliefs to explain a phenomenon rather than use the data at hand. In other cases, students may use data from an investigation, but not the appropriate data. If students are confronted with more data than is appropriate to use as evidence for a particular claim, they can have a difficult time differentiating between what is appropriate and what is inappropriate (McNeill & Krajcik, in press). Besides appropriateness, students can also struggle with providing sufficient evidence. Students can realize the importance of including data in their explanations, yet they often still do not include sufficient evidence for their claims (Sandoval & Millwood, 2005). They may only rely on one piece of evidence when multiple pieces may be necessary to construct a strong explanation.

Students' understanding of the content can also influence their ability to effectively use evidence in their explanations. Students are more likely to discount data if the data contradicts their current theory (Chinn & Brewer, 2001) and they are more likely to consider data if they can come up with a mechanism for the pattern of data (Koslowski, 1996). More general reasoning strategies interact with domain-specific knowledge when students evaluate data (Chinn & Brewer, 2001). This suggests that both students' understanding of the content and an understanding of evidence can influence students' ability to provide evidence for a particular task. Whether students provide appropriate and sufficient evidence may depend on students' understanding of what counts as evidence, the particular task, and students' content knowledge (McNeill & Krajcik, in press).

Reasoning. Students also have difficulty providing the backing, or what we refer to as reasoning, for why they chose the evidence in their written explanations (Bell & Linn, 2000).

Reasoning is the logic for why the evidence supports the claim, which can often include scientific principles. Other researchers have shown that during classroom discourse, discussions tend to be dominated by claims with little backing to support their claims (Jiménez-Aleixandre, Rodríguez & Duschl, 2000). In our previous work, we found that middle school students had the most difficulty with the reasoning component of scientific explanations. Although students often linked their claim and evidence, they less frequently articulated the scientific principles that allowed them to make that connection (Lizotte, Harris, McNeill, Marx & Krajcik, 2003; McNeill, Lizotte, Harris, Scott, Krajcik & Marx, 2003). This is similar to the findings of other research that shows that individuals often do not explicitly use canonical scientific knowledge in science-related everyday situations (i.e. a parent's choice to immunize a child) or practical science related professions (i.e. nursing) in order to support their decisions (Aikenhead, 2004).

Similar to students' ability to evaluate and use data, providing accurate reasoning may also be related to students' understanding of the content. Previous research with students has found that their success at completing inquiry practices is highly dependent on their understanding of both the content and the inquiry practices (Metz, 2000). If students do not understand the scientific principles, they will not be able to apply those principles to a particular scientific inquiry practice, such as constructing explanations. Yet understanding the science content alone may not be sufficient for providing appropriate reasoning in a scientific explanation.

The numerous studies on scientific explanation and argumentation suggest that this ability does not come naturally to most individuals, but rather is acquired through practice (Osborne, Erduran & Simon, 2004). Students need to be explicitly taught about scientific explanation in order to be successful. We focus on how to support students' construction of scientific explanations in classrooms in order to better prepare students to write explanations when they are not provided with support.

Our Instructional Model for Scientific Explanation

In our work, we chose to use the phrase "scientific explanation" with teachers and students in order to be consistent with science standards (American Association for the Advancement of Science, 1993; National Research Council, 1996). The usability of an education innovation is determined in part by its alignment with school culture (Blumenfeld, Fishman, Krajcik, Marx & Soloway, 2000). Traditionally, there has been a gap between educational research and classroom practice. In order to help bridge that gap, there needs to be more research-based development of tools and models for practitioners that is better linked to the practical needs of the education system (Burkhardt & Schoenfeld, 2003). Currently, national standards and high stakes testing play a large role in school culture. In order to create a more usable and sustained innovation, we feel it is important to align our model of scientific explanation with national science standards. Both the American Association for the Advancement of Science's (AAAS) *Benchmarks* (1993) and the National Research Council's *National Science Education Standards* (1996) discuss the importance of creating and critiquing explanations with appropriate evidence and reasoning. For example, one of the AAAS middle school scientific inquiry standards states, "…scientific investigations usually involve the collection of relevant evidence, the use of logical reasoning, and the application of imagination in devising hypotheses and explanations to make sense of the collected evidence." (1993, p. 12). Repeatedly, the NRC's inquiry abilities stress the importance of developing explanations using evidence. In the NRC's understandings about scientific inquiry they state "…Scientists evaluate the explanations proposed by other scientists by examining evidence, comparing evidence, identifying faulty reasoning, pointing out statements that go beyond the evidence, and suggesting alternative explanations for the same observations." (1996, p. 148). Our goal is to create an instructional model of scientific explanation that is usable by a large number of teachers and students. Consequently, we have aligned our model of scientific explanation with the national standards though our work is informed by research both on explanation and argumentation.

Many science educators have used Toulmin's model of argumentation to support students in both explanation and argumentation (Bell & Linn, 2000; Driver, et al., 2000; Erduran, Simon & Osborne, 2004; Jiménez-Aleixandre, Rodríguez, & Duschl, 2000; Sandoval, 2003; Zembal-Saul, et al., 2002). We decided to adapt Toulmin's model instead of using it directly in order to more closely align with the standards and to create an instructional model that we thought might be more accessible to middle school teachers and students. As Toulmin's model has been difficult to interpret at times for researchers and philosophers (van Eemeren et al., 1996), we felt that it would be difficult for middle school students to use.

We view our instructional model of scientific explanation as an entry point into this important scientific inquiry practice for middle school students. Creating explanations is a complex task for students. Our main objective is to reduce the complexity for students (Quintana et al., 2004) and to focus the attention of the learner on the relevant task features (Pea, 2004) so they can more easily accomplish creating scientific explanations. As a result, we created an instructional model that breaks down explanation into three components: a claim (similar to Toulmin's claim), evidence (similar to Toulmin's data), and reasoning (a combination of Toulmin's warrant and backing). The claim is an assertion or conclusion that answers the original question. The evidence is scientific data that supports the claim. These data can come from an investigation or from another source, such as observations, reading material, or archived data. The data need to be both appropriate and sufficient to support the claim. The reasoning is a justification that shows why the data count as evidence to support the claim. In the reasoning component, we encourage students to articulate the logic behind why they believe the evidence supports the claim, similar to Toulmin's warrant. Furthermore, students may need to back up that link between the claim and evidence by including the appropriate scientific principles, similar to Toulmin's backing. We decided to use the term, "reasoning", in order to limit the number of components and to use a term that was in the standards and more familiar to both teachers and students.

Use of our instructional model across content and contexts. We developed our instructional model of scientific explanation to be used across different content and in different contexts in middle school science curriculum. Yet this is a contentious issue. Whether or not teaching general strategic knowledge proves useful for reasoning in context is a complicated and unresolved issue (Perkins & Salomon, 1989). There are practices, which are common across domains. For example, argument is an example of a general cognitive practice that transcends the particular content to which it refers (Kuhn, 1993) and is an essential critical thinking skill across different age levels and subject areas (Yeh, 2001). Argument has been a goal across a variety of domains (Voss & Van Dyke, 2001) including language arts (Reznitskaya & Anderson, 2002; Reznitskaya, Anderson, McNurlen, Nguyen-Jahiel, Archodidou, & Kim, 2001), economics (Cho & Jonassen, 2002), mathematics (Cobb, 2002), debate (Kuhn & Udell, 2003), and science (Driver, et al., 2000). The structure of the argument has some similarities across the domains, though the content and context are also important.

Specifically in science, there are scientific reasoning abilities or scientific inquiry practices that transcends specific contexts. A key goal for science education is to help students seek evidence and reasons for the ideas or knowledge claims that we draw in science (Driver et. al., 2000). Scientists across domains value supporting ideas with evidence. Yet both content knowledge and scientific reasoning skills are important for success at a particular task (Kuhn, Schauble, & Garcia-Mila, 1992). Domain specific knowledge determines the types of questions asked, the methods used, and what counts as evidence (Passmore & Stewart, 2002; Sandoval, 2003). For example even in two areas of chemistry, a synthetic chemist who prepares new compounds relies on multiple sources of spectroscopic data as evidence to deduce possible molecular structures, while a theoretical chemist building a model for how proteins fold relies on how well the computer output on the molecular structure matches an experimentally determined structure. Considering the content and context is necessary for determining the appropriateness and strength of the explanation, not just the structure alone.

Although content knowledge is essential, using the same explanation instructional model across domains may help students become more adept at writing explanations. We acknowledge that using this general framework does simplify the complex task of constructing an explanation and open the possibility of misrepresenting it. Yet, using a generic framework across different science content areas and contexts, can help students achieve a basic understanding of the processes and practices of science and help students better understand how knowledge claims are created and supported (Osborne, Collins, Ratcliffe, Millar & Duschl, 2003). We conjecture that our scientific explanation model of claim, evidence, and reasoning is an appropriate instructional model for students in the middle grades. In this study, we explore how to best support middle school students' use of our model of scientific explanation.

Scaffolding Student Learning

In order to help middle school students learn how to construct scientific explanations, we incorporated our instructional model into an 8-week project-based chemistry unit. Research on scaffolding informed both our design of the curriculum materials and this research study where we investigated the effect of fading written instructional supports (scaffolds) for scientific explanation.

Wood, Bruner and Ross (1976) originally introduced the term “scaffolding” in the context of adult-child interactions where the more knowledgeable adult tutors the child to complete a task the child would be unable to do on his/her own. With the help of scaffolds, learners can complete more advanced activities and engage in more advanced thinking (Bransford et al., 2000). Although Wood et al. did not originally connect scaffolding to Vygotsky’s zone of proximal development, a number of educational researchers since then have explicitly made this connection (Hogan & Pressley, 1997; Palincsar & Brown, 1984). The zone of proximal development (ZPD) defines the area between a child’s independent problem solving capabilities and the level of potential problem solving capabilities with the guidance of people or tools (Vygotsky, 1978). Stone (1993) argues that scaffolds allow students to achieve a higher level of understanding within their zone of proximal development. In order for a scaffold to promote student understanding, it needs to reside within a students’ current ZPD. If a scaffold provides too much information, the student will not be challenged to learn more. The scaffold should provide just enough information that the learner may make progress on his/her own (Hogan & Pressley, 1997).

In their study of reciprocal teaching, Palincsar and Brown (1984) discuss Vygotsky's idea that at first the parent or expert guides much of a child's cognitive activities and over time the child takes on more and more of those responsibilities. Eventually, the child performs the activities by herself, without the help of the scaffolds. In fact, Wood et al. (1976) described scaffolding as a flexible process contingent on what a child knows and the characteristics of the learning task. This suggests that scaffolds should be adjusted over time rather than remaining constant in order to allow students greater responsibility over their own learning. Palincsar and Brown's study (1984) supports this idea of adjusting scaffolds based on students' understanding. In studying teacher-student interactions during reciprocal teaching, they found that initially the teacher provided modeling, feedback, and practice to students. Over time as the student became better able to complete a task, the teacher decreased his or her support. By the end, the teacher's role was one of supportive audience member and the student had taken over the expert responsibilities. This shift to greater control over knowledge construction resembles the shift from child to adult status where adults retain a more regulatory role controlling the cognitive interaction in their ZPD (Scardmalia & Bereiter, 1991).

Decreasing the support or "fading" (Collin, Brown & Newman, 1989) is an essential characteristic of scaffolds. There are numerous supports within learning environments that are not scaffolds; rather they can be instrumental for both experts and novices to complete a task. In these situations, the supports serve to distribute the knowledge across the physical environments (Brown, Collins, & Duguid, 1989). Such supports have been given various names such as acts of distributed cognition or distributed intelligence (Pea, 2004) and cultural tools (Tabak, 2004). Tabak (2004) describes the difference between a cultural tool and a scaffold by using the example of a vertical array representation to find the solution to a multiplication problem, such as 343 multiplied by 822. Both experts and novices might use the vertical array to find the solution; consequently, it is considered a cultural tool. In contrast, scaffolds serve to help learners complete a task independently and as such should be faded as learners develop their own understanding. For example, a novice unfamiliar with using a vertical array may be provided with a special notational template to help them complete multiplication problems. Eventually, when learners achieve mastery of the vertical array representation, the template would no longer be used, so it is considered a scaffold. Based on this idea that scaffolds are specifically developed to help learners, we define scaffolds as temporary supporting structures provided by people or tools to promote learning of complex problem solving.

Traditionally, scaffolding has been discussed in terms of one-on-one interactions. There has been little research on teacher-student scaffolding in whole class settings (Hogan & Pressley, 1997). Hogan and Pressley argue that one of the reasons there has been little research in this area is because in a large classroom a teacher cannot possibly interact with every child individually. Ideally, the teacher would react to the current situation and modify the scaffolds based on all of the students' needs. When a teacher addresses the whole class he or she is confronted with multiple zones of proximal development. There is concern that teacher-student scaffolding cannot be carried out effectively in such whole class settings (Stone, 1998).

One possible solution to this problem is having students work in groups and then scaffolding those groups. But this can still be problematic because of the number of groups in a classroom. Another possibility is to provide students with tools, such as computers or written materials, which provide students with scaffolds. Here the interaction is between the student and the computer or written materials. Because external tools (like computers or written artifacts) cannot include the dynamics of adult-child or even peer interactions, they can be seen as limited in the use of the scaffolding metaphor (Stone, 1998). Palincsar argues that one way researchers "have hobbled the use of scaffolding is by attributing scaffolding only to interactions that occur between individuals, and typically between individuals of significantly different expertise…(I)t is helpful to recall that ZPDs include not only people but also artifacts, and that ZPDs are embedded in activities and contexts" (1998, p. 371).

Fading Written Scaffolds

Although previous research suggests fading encourages greater student independence, the majority of these studies have looked at adult-child interactions where the adult can individualize the scaffolds for the particular student's needs. Written supports obviously do not have that advantage though continuous written prompts, which are provided throughout an instructional unit, have been shown to increase student learning. One example of the benefits of continuous written prompts includes the ThinkerTools curriculum created by White and Frederiksen (1998; 2000). They designed their curriculum to scaffold students' development of scientific inquiry processes, modeling, and metacognitive skills and knowledge. In order to develop metacognitive skills, they developed a set of reflection prompts that guided students' evaluation of their work at the end of each phase of the inquiry cycle. To determine the effectiveness of the metacognition prompts, White and Frederiksen compared two versions of the curriculum, one with reflection prompts and one without reflection prompts. They found that students who received the reflective prompts resulted in greater understanding of the inquiry practices.

Davis (2003) also examined the role of continuous prompts in supporting students' reflection. In this case, she integrated the prompts into the Knowledge Integration Environment (KIE) software where she investigated the role of two different types of reflection prompts:

generic prompts and directed prompts. She found that generic prompts were more productive for student reflection than directed prompts. Davis argues that the generic prompts were more effective because they allowed students to partially define the support and take more control over their own reflection. Her study suggests that written instructional prompts that encourage greater responsibility, such as scaffolds that fade to provide less support, may be more productive for student learning in classroom contexts that are heavily scaffolded in other ways.

In terms of scientific explanation and argumentation, recent research suggests that computer software and written supports can help students complete this complex task. For example, Sandoval (2003) found that providing students with prompts that provide domain-specific guidance through a software program, ExplanationConstructor, helped students produce useful explanations for their inquiry. The explanation supports highlighted the causal components of important domain theories for the particular task, using natural selection to explain the population change of finches on the Galapagos Island of Daphne Major. The supports appeared to help focus students on the explanatory goals of the task. Bell and Linn (2000) investigated the impact of an argument building software tool called SenseMaker on middle school students' construction of arguments about light propagation. They found that the arguments students constructed did link their evidence with their conclusions or claims and promoted knowledge integration. In terms of written prompts to support explanation, Lee and Songer (2004) found that providing 5^{th} and 6^{th} grade students with written prompts for explanations during a biodiversity unit resulted in students having a stronger understanding of the content and increased their ability to match given evidence to their claims.

There is a growing body of research exploring scaffolded tools (Quintana et al, 2004), such as Computer Supported Intentional Learning Environments or CSILE (Scardamalia &

Bereiter, 1991), SMILE (Kolodner & Nagel, 1999; Owensby & Kolodner, 2004), Artemis (Krajcik, Blumenfeld, Marx & Soloway, 2000), WorldWatcher (Edelson, Gordon & Pea, 1999), and The Galapagos Finches (Reiser et al, 2001). This work considers classrooms as highly complex systems where student learning can be mediated in many ways including the support of technological tools (Davis & Miyake, 2004). Our study builds off of this work as well as scaffolding research on written prompts and adult-child interactions. Although the research community has taken great strides in its understanding and design of scaffolds, many questions remain including the issue of fading scaffolds (Davis & Miyake, 2004). We address the question of whether you can fade written supports when there is no individualization afforded such as in adult-child interactions, specifically, for students' construction of scientific explanations.

Designing Our Explanation Scaffolds

In scaffolding students' explanation construction, we attempted to make our explanation framework clear to students (and teachers) in order to facilitate their understanding of what an explanation is and how to construct one. Making scientific thinking strategies explicit to students can facilitate their use and understanding of these strategies (Herrenkohl, Palincsar, DeWater, & Kawasaki, 1999). More specifically, revealing the tacit framework of scientific explanation through scaffolds can facilitate students' explanation construction (Reiser et al, 2001). If students do not initially provide their reasoning, prompting can result in students articulating their thoughts about how and why something occurs (Chinn & Brown, 2000). We hoped that by providing students with our explanation framework, we would encourage deeper thinking and promote students translation of their thinking into written text.

Combining generality and context specificity in instruction can result in greater student understanding and ability to use cognitive skills (Perkins & Salomon, 1989), such as constructing scientific explanations. Since both an understanding of the content and scientific explanation are important for the creation of explanations, we created written supports that included both generic and context-specific support. Our prompts included generic support for claim, evidence and reasoning that we repeated regardless of task. Generic support helps students understand a general framework for their explanation regardless of the content area. The prompts also included context specific support that changed with the task. Context specific support provides students with hints about the task and what content knowledge to use or incorporate into their explanation. Figure 1 displays an explanation prompt from the unit where the generic portion is circled. For example, the generic support of the claim prompt is "**Claim** (Write a sentence that states)". This prompt is repeated in every case where the student is provided a complete claim scaffold. The next part of the prompt is context specific, "(whether the nail and wrench are the same or different substances.)" This prompt is specific to the task, which is about a nail and a wrench, and it is specific to the content, which is about substances and properties. We conjectured that by using this repeated format that students would understand how this same explanation framework could be used and adapted across multiple contexts and content areas.

Yet we still wondered whether fading the written prompts would be effective. Scaffolds should be sensitive to students' current understanding and provide just enough information that students can precede on their own (Hogan & Pressley, 1997). As students begin to learn the explanation framework, the prompts should be adjusted or faded to students' current understanding. Fading scaffolds can make learning tasks more difficult for students in the short term, but ultimately promotes student learning (Reiser, 2004). When fading scaffolds, this forces students to think about what they have learned from the previous scaffolds and apply their knowledge to the current learning task. The removal of the scaffolds is our ultimate goal since we want students to be able to complete the task independently. We conjecture that the fading of written prompts may help students achieve this goal of independence by providing a more challenging situation that encourages them to think through the different components of explanation. But the danger of fading a written support is that since it is not individualized it may fade too quickly and reside outside of a child's ZPD. Previous research providing 5th and 6th grade students with context specific written supports found that fading prompts resulted in less student learning (Lee & Songer, 2004). However, we conjecture that with older students, repetitive prompts that contain both generic and context specific portions could fade to produce greater student learning.

We created two written prompt treatments: Continuous and Faded. The Continuous group received the same type of written prompt at six time points during the unit on their investigation sheets. The prompts provided detailed information about each explanation component. The Faded group also received written prompts on the same six investigation sheets, but these supports provided less detail over three stages each of which included two written scaffolds (see Table 1). We realize that the fading of scaffolds should be purposeful. We chose equal intervals for each stage, because we did not have a reason to provide students with one of the prompts for a longer amount of time than the others.

For the Faded group, the written prompts in Stage I was identical to the Continuous group, while Stage II and Stage III provided increasingly less support. For example, during Stage I for evidence the written prompt stated, "TwoPieces of Evidence (Provide twopieces of data that support your claim that new substances were or were not formed.)" followed by two written prompts that said “Evidence #1” and “Evidence #2.” The Continuous group received evidence prompts similar to this one on all six investigation sheets. The Faded group received the same detailed prompt as the Continuous group only in Stage I. During Stage II, the Faded group received an intermediate prompt for evidence, which stated “Evidence (Provide data that support your claim.).” Finally, during Stage III the Faded group received the least supportive prompt, which simply stated, “Remember to include claim, evidence, and reasoning,” with no specific prompts about the different components. Table 2 provides a timeline of when during the unit students wrote explanations, the stage of the scaffold for the different lessons, and a description of the specific learning task.

Introduction of Explanation

Although we wanted to investigate the role of written explanation scaffolds, we decided that the written prompts were not a sufficient introduction to help middle school students create scientific explanations. In order for scaffolding to be successful, a child must have some prior understanding of what is to be accomplished (Wood, Bruner, & Ross, 1976; Stone, 1998). This is important because it means in a complex learning environment an inquiry practice, like constructing explanations, should not be introduced through written scaffolds. Instead, the teacher needs to first help students understand the inquiry practice before they can effectively use the written scaffolds embedded in the curriculum. For example, Chen and Klahr (1999) found that providing students with the rationale behind controlling variables in science experiments, as well as examples of unconfounded comparisons before completing investigations, resulted in greater learning relative to students who did not receive the explicit instruction. Developing a deep understanding of scientific inquiry practices may require a variety of different material and social supports, each of which offers different affordances and constraints that work in concert over time (Tabak, 2004).

One of the essential social supports is the teacher, who is important in the successful use of a scaffolded tool (Pea, 2004). In our unit, the teacher introduces the explanation framework and models how to construct explanations during a focal lesson before students receive the written scaffolds. Similar to the written scaffolds, this lesson combines both context specific and generic support. The teacher introduces scientific explanation in the context of whether fat and soap are the same or different substance. Initially, the teacher provides the general explanation framework, but then models and critiques multiple examples of explanations in the context of fat and soap. This allows students to obtain an initial understanding of scientific explanations so later when they are provided with the written scaffolds they have some prior understanding of what they need to accomplish.

The Relationship Between Explanation and Science Content

In our analysis of students' explanations, we examine both their ability to construct explanations and their understanding of the science content. In order to construct an accurate scientific explanation, students need to understand both the content and how to construct a scientific explanation. Metz (2000) argues that, "the adequacy of individuals' reasoning is strongly impacted by the adequacy of their knowledge of the domain within which the reasoning is tested. Thus, inside the research laboratory and beyond, cognitive performance is always a complex interaction of scientific reasoning capacities and domain-specific knowledge" (p. 373). If students have difficulty with any of those components, they will be unable to write an accurate scientific explanation. For example, if a student does not understand the content even though they understand how to write an explanation, they will be unable to construct an accurate explanation. Consequently, we need to look beyond one explanation to hypothesize why students may be having difficulty with explanations.

The present paper examines results of an enactment of our unit in which students in different classes received one of two written prompt treatments: Continuous, involving detailed support for every investigation, or Faded, involving less support over time. We address the following research question: Do the two written prompt treatments have different effects on students' explanations throughout the unit?

Method

Instructional Context

Using a learning-goals-driven design model (Reiser, Krajcik, Moje, & Marx, 2003), we developed a middle school chemistry unit (McNeill et al., 2003) as part of the *Investigating and Questioning our World through Science and Technology* (IQWST) curriculum materials. Learning-goals-driven design uses key learning goals identified from the national standards (American Association for the Advancement of Science, 1993; National Research Council, 1996) to guide all phases of curriculum and assessment design. The IQWST curriculum materials are currently being developed in a collaborative effort by researchers at the University of Michigan, Northwestern University, the University of Illinois Urbana-Champaign, Michigan State University, and Columbia University. We used this design model to develop an 8-week project based unit addressing the driving question (Krajcik, Czerniak, & Berger, 1999), "How can I make new stuff from old stuff?" (McNeill, Harris, Heitzman, Lizotte, Sutherland &

Krajcik, 2004). The unit is grounded in whether you can make soap from fat. During the instructional sequence, students completed fifteen lessons where they investigate a variety of phenomena, yet they repeatedly cycled back to soap and fat. Each cycle helped them delve deeper into the science content initially to understand substances, then properties, chemical reactions, and finally the conservation of mass at both the macroscopic level and in terms of the particulate nature of matter. We designed the unit around a number of key design principles including a focus on having students experience multiple and varied phenomena as well as the importance of dialogue and collaboration among students during these experiences (Reiser et al., 2003).

Besides the science content, our other key learning goals focused on scientific inquiry practices such as the construction of scientific explanations. As mentioned earlier, in order to introduce students to scientific explanations, we developed a one day focal lesson. Approximately two weeks into the unit, students gathered all of the data they had collected on fat and soap such as density, melting point, color, solubility and hardness. The investigation asked students “Write a scientific explanation stating whether you think fat and soap are the same substance or different substances.” The students did not receive written prompts for this scientific explanation. First students wrote explanations using their own data and their prior understanding of scientific explanations. Then the teacher led a discussion about scientific explanation in order to make the instructional model (claim, evidence, and reasoning) explicit to students. The teacher also modeled the construction of scientific explanations through the use of hypothetical examples of weak and strong explanations. Using this framework and models, students revised their original explanations.

After the focal lesson, students wrote a number of explanations during the unit to explain the results for both first hand investigations where they collected the data and for second hand investigations where they were provided with data (see Table 2). Typically, an investigation took one or two class periods and students wrote their explanations at the end of the investigation after they had analyzed the data. Students recorded scientific explanations on student investigation sheets. Each student had their own lab book, which contained his or her investigation sheets. As we mentioned previously, we created two scaffold treatments, Continuous and Faded. Students' lab books contained one of the two types of written prompts (see Table 1). Students worked in groups to collaboratively complete the investigation, collect the data, and discuss the phenomena. Yet each student was responsible for writing his or her own explanation. Students were supported in writing their explanations by both the written prompts on their investigation sheets as well as other social supports, such as their peers and teachers.

Participants

Participants included 6 teachers and 331 7th grade students from schools in the Mid-west. Three of the teachers and 260 of the students in 9 classes were from public middle schools in a large urban area. The majority of these students were African American and from lower to lower-middle income families. The other three teachers and 71 students in 5 classes were from an independent middle school in a large college town. The majority of these students were Caucasian and from middle to upper-middle income families.

Data Sources

We collected two types of assessment data: student investigation sheets and pretest and posttest data. We scored all student explanations by adapting our base explanation rubric (Appendix A). A base rubric is a general rubric for scoring an inquiry practice across different content and learning tasks. The base explanation rubric includes three components (claim, evidence, and reasoning) and the scoring levels for each component (Lizotte et al., 2003). Base rubrics can be adapted for any science content at the middle school level. We adapted the base rubrics to create specific rubrics with tailored scoring levels. Sandoval and Millwood (2005) argue for the importance of assessing the conceptual adequacy of students' arguments along with structural analyses. Our method of adapting a basic explanation rubric to a specific content area and tasks combines both structure and content. Explanations that receive the highest score include accurate science content and the appropriate explanation structure to support the claim.

The different scoring levels of the specific rubric depend on the content, such as substance/ property versus chemical reaction, and the specific learning task, such as the chemical reaction during electrolysis compared to making soap. Appendix B provides an example of a specific rubric for open-ended test question #1, which asks students whether any of four solids are the same substance. A more complete description of our coding process and examples of student work can be found in McNeill and Krajcik (in press). By adapting the same base rubric, we can compare students' scores for the different explanation components across the different content areas to see if students are providing appropriate and sufficient evidence and reasoning.

For the student investigation sheets, all three components of explanations (claim, evidence, and reasoning) were scored separately using appropriate specific rubrics. Again, the sufficiency of the components varied by task, but in all cases an appropriate and complete response received the highest score. All questions were scored by one rater. We then randomly sampled 20% of the student sheets and a second independent rater scored them. For each of the seven explanations our estimates of inter-rater reliability were calculated by percent agreements. Our inter-rater agreement was above 96% for claim, 88% for evidence, and 85% for reasoning on each explanation.

All students completed the same pretest and posttest, which consisted of thirty multiple-choice and eight open-ended items. The total possible score on the test was 60 points with 30 points for the multiple-choice items and 30 points for the open-ended items. We weighted each open-ended item so that all eight items had an equal value of 3.75 points. The test required two class periods to complete, which occurred over two consecutive days. The test was taken the two school days before beginning the unit and the two school days after finishing the unit. Only students who completed all parts of the test were included in the analysis. Due to high absenteeism in the urban schools and the necessity of students being in class for all four days of testing, only 220 students took all parts of the pre- and posttest assessments. We randomly selected 20% of the pretests for students that we did not have posttests. In our missing data analysis, we only examined the multiple-choice portion of the test since the majority of students missing complete exams were excluded because they were missing the open-ended items. These students' multiple-choice scores ($M = 13.2$, $SD = 3.8$) did not significantly vary from those of the other students who had complete test data ($M = 12.5$, $SD = 4.1$) in their respective classes, $t\ (172) = 0.901$, *ns*.

For this study, we focused on the four open-ended items, which asked the students to write a "scientific explanation" (see Appendix B for examples). Two of the items focused on the substance and property content, while the other two items addressed chemical reactions. Each student received all four items, two items had the detailed Continuous-type scaffolds and the other two items had no scaffolds. We created two versions of the test to counterbalance which items had scaffolds and no scaffolds across students in order to ensure that any differential effects were not the result of varying question difficulty. For all four questions, we scored the different components of explanation (claim, evidence, and reasoning) separately using appropriate specific rubrics.[1] All items were scored by one rater. Twenty percent of the tests were randomly chosen and scored by a second independent rater. The inter-rater agreement for the two scaffolded items was above 94% for claim, 86% for evidence, and 85% for reasoning. The inter-rater agreement for the two unscaffolded items was above 96% for claim, 90% for evidence, and 90% for reasoning.

Study Design

We randomly assigned classes of students to the Continuous and Faded groups so that teachers with multiple classes taught both groups. For example, if a teacher taught two classes of seventh grade science, we assigned one class the Continuous treatment and the other class the Faded treatment. We charted students' explanations through successive stages of the unit. Stages I, II, and III occurred sequentially during the full 8-week enactment and each involved two investigation sheets for explanation (see Table 2). For students in the Continuous scaffold treatment, scaffolds on the investigation sheets were identical in Stages I, II, and III, whereas students in the Faded scaffold treatment received progressively less detailed scaffolds through Stages I, II, and III.

Of the 220 students who completed both the pre- and posttest, we were only able to collect the student investigation books for a subset of those students. One of the challenges of working in numerous classrooms with teachers who volunteer their time is the collection of student artifacts. We obtained student investigation books that included a completed focus lesson explanation and at least one explanation for each of Stages I, II, and III for 129 of the students. We charted explanations of those 129 students through the unit. However, we used the larger sample of students for all quantitative analyses of pre and posttest data to provide us with more power in detecting any significant effects.

Results

Our analyses address three specific questions: 1) How do the different scaffold treatments (Continuous or Faded) during the unit influence students performance on explanations they write in class? 2) Do the scaffold treatments during the unit have different effects on students' performance on explanations on posttest items with and without scaffolds? and 3) Does the influence of scaffolds vary depending on the content and component of the explanations? These questions allow us to address our larger question of whether fading written instructional supports influences student learning of scientific explanation. By examining students' explanations both during the unit where they are provided with written scaffolds and on their pretests and posttests, we can see whether the effect of the scaffolds varies depending on whether students have other supports available to them. As we mentioned earlier, during the unit each student was responsible for constructing their own explanation though this occurred within a complex learning environment where other social supports, such as peers and the teacher, were available. On the pre and posttest, students worked independently without other supports. We are interested in the effect of the scaffolds in both situations. We also wondered if the effect varied by explanation component or content since previous research suggests that certain aspects (such as reasoning) are more difficult for students.

Influence of Scaffolds During the Unit on Students' In Class Explanations

We are interested in whether the written prompts students received during the unit influenced the explanations students wrote in class. Figures 2, 3, and 4 chart the mean scores for claims, evidence, and reasoning, respectively, through the stages of the unit for students in the Continuous and Faded treatments from whom we received student investigation books. Interestingly, the three figures show an increase in students' claim, evidence and reasoning scores regardless of treatment from the pretest to the focal lesson where the teacher explicitly introduced students to our model of scientific explanation. To test whether there were differences in students' explanations for the Continuous and Faded groups during the unit, we performed separate repeated measures ANOVAs on claim, evidence, and reasoning scores, using a multivariate approach. Each ANOVA was 2 x 6, with Scaffold Treatment (continuous, faded) as the fixed factor and Time (pretest, focal lesson, Stage I, Stage II, Stage III, posttest) as the repeated factor. For the pre- and posttest, we collapsed students' scores across test items with and without scaffolds.

For the analysis of students' claims, there was a significant main effect of Time, F (5, 123) = 77.82, $p < .001$, but there was no main effect of Scaffold Treatment. Students' evidence scores showed a similar result where there was significant main effect of Time, F (5, 123) = 69.99, $p < .001$, but no main effect of Scaffold Treatment. Finally for reasoning, students' scores again showed a significant main effect of Time, F (5, 123) = 61.77, $p < .001$, but no main effect for the scaffold treatment. This suggests that students' performance is changing over time, as we will discuss below in more detail for students pre and posttests. Yet during the unit when students are in a complex classroom environment where multiple supports are available to them, there is not a significant difference in students' written explanations in the two treatments.

Influence of Scaffold Treatments on Posttest Explanations

We examined whether the Faded and Continuous scaffolds treatments influenced students' explanations on the test items using the entire sample of 220 students (n = 97 for continuous treatment; n = 123 for faded treatment). Students in both the Faded and Continuous groups had significant pre-posttest gains during the unit on all three components of explanation for the four scientific explanations (Table 3). We tested whether a scaffold treatment effect was present for test items with scaffolds, without scaffolds, or both types, by performing separate ANCOVAs on students' posttest claim, evidence, reasoning, and composite scores for items with and without scaffolds. For each ANCOVA, Scaffold Treatment (continuous, faded) was the fixed factor and the appropriate pretest score was the covariate. The effect of the pretest covariate was significant in each analysis.[2] The effect of Scaffold Treatment was marginally significant in one analysis: reasoning scores on posttest items without scaffolds were higher for students in the Faded treatment than the Continuous treatment F (1, 217) = 3.28, p = .07. Figure 5 shows the mean reasoning scores for posttest items with and without scaffolds. This finding suggests that fading written scaffolds during instruction can produce greater student gains for the reasoning component of explanation for items without scaffolds.

To further tease apart this effect on reasoning, we examined the substance/property and chemical reaction posttest explanations separately to evaluate the role of science content. We found that this effect of scaffold treatments on reasoning scores for test items without scaffolds applied to explanations about substance/property phenomena, but not chemical reaction phenomena. We performed separate ANCOVAs on students' posttest reasoning scores for substance/property and chemical reaction items, with scaffolds and without scaffolds; Scaffold Treatment (continuous, faded) was the fixed factor and the appropriate pretest score was a covariate for each analysis. The effect of the pretest covariate was significant in all analyses except for the analysis of reasoning scores on chemical reaction items with scaffolds, $F\ (1, 217) = 1.82$, *ns*.

For substance/property items, the effect of Scaffold Treatment on students' reasoning scores was significant for items without scaffolds, $F\ (1, 217) = 3.95$, $p < .05$. Figure 6 shows that the mean posttest reasoning score for substance/property items without scaffolds was higher for students in the Faded treatment than the Continuous treatment. For comparison, Figure 6 also includes the mean posttest reasoning scores for substance/property items with scaffolds. Again, the Faded treatment is higher, but this difference is not significant. Fading written scaffolds during instruction had a positive effect for items without scaffolds on the posttest. For test items that included scaffolds, the type of scaffolding treatment during the unit was not significant.

For chemical reaction items, there were no significant effects of Scaffold Treatment on students' reasoning scores. Specifically, posttest reasoning scores for students in the Faded and Continuous treatments did not differ either for chemical reaction items without scaffolds ($M = 0.39$, $SE = 0.04$ for Faded; $M = 0.33$, $SE = 0.04$ for Continuous) or for those with scaffolds ($M = 0.40$, $SE = 0.04$ for Faded; $M = 0.35$, $SE = 0.04$ for Continuous).

Considering that we only observed a significant difference on test items without scaffolds, we were interested in whether the presence or absence of scaffolds on the posttest influenced the students in the Continuous treatment differently than the students in the Faded treatment. Figures 7 and 8 show students' reasoning scores on substance/property items both with and without scaffolds. We performed separate repeated measures ANCOVAs on students' posttest reasoning scores for substance/property items for the faded and continuous groups. Test Item Scaffold (scaffold, no scaffold) was the repeated factor and the appropriate pretest score was a covariate for each analysis in order to control for students' scores at the beginning of the unit. The effect of the pretest covariate was significant for the Continuous group; it was marginally significant for the Faded group, $F\ (1, 121) = 3.40, p < .07$. For neither group did the effect of the covariate differ for items with and without scaffolds, $F\ (1, 95) = 0.88$, *ns* for Continuous, $F\ (1, 121) = 1.57$, *ns* for Faded.

Results indicated no effect of Test Item Scaffold on the posttest reasoning of students in the Faded group, $F\ (1, 121) = 2.63$, *ns*. Figure 7 shows that before the unit for the Faded group, students' scores on the scaffolded and nonscaffolded items were farther apart than after the unit. This suggests that the Faded treatment helped students improve at constructing explanations without scaffolds. The scaffolds no longer played an important role on the posttest for their reasoning. For the Continuous group, there was an effect of Test Item Scaffold on their posttest reasoning, $F\ (1, 95) = 4.30, p < .05$. Students in the Continuous group had higher reasoning on the posttest items with scaffolds than those items without scaffolds. Figure 8 shows that the difference between the Continuous students' scores on items with and without scaffolds became greater over the unit. This suggests that the Continuous students still depended on the scaffolds at the end of the unit and had a more difficult time constructing explanations without scaffolds than with scaffolds.

To summarize, students who received the Faded treatment had significantly higher reasoning scores on posttest items without scaffolds than those who received the Continuous treatment, but only for substance/property items. For the students in the Faded treatment, there was also not a significant difference between their posttest reasoning scores for items with scaffolds compared to those without scaffolds, while there was a significant difference for students in the Continuous treatment. The students in the Continuous group constructed stronger explanations on the posttest when they were provided with the written scaffolds. These results suggest that the scaffold treatment had different effects on the different components of explanation (claim, evidence, and reasoning) and the different content areas (substance/property and chemical reactions).

Relationship Between Science Content and Scientific Explanations

To further investigate this relationship between the science content and students' ability to construct explanations, we looked at students' performance on the multiple-choice items for both substance/property and chemical reaction. We determined the correlations between students' posttest multiple-choice and explanation scores for each content area. Not surprisingly, there is a relationship between these two scores. Students' scores on the substance/property multiple-choice items were significantly correlated with their substance/property explanations, *r*s (220) = 0.37 for claim, 0.35 for evidence, and 0.52 for reasoning, *p*s < .001. Students' scores on the chemical reaction multiple-choice items were significantly correlated with their chemical reaction explanations, *r*s (220) = 0.45 for claim, 0.40 for evidence, and 0.41 for reasoning, *p*s < .001. Students who had higher multiple-choice scores in a content area also had higher explanation scores in that area.

We examined students' posttest scores comparing substance/property items to the chemical reaction items. We found that students scored higher on the set of substance/property items (*M* = 3.90, *SE* = 0.07) than the equally-weighted[3] set of chemical reaction items (*M* = 3.65, *SE* = 0.08), t (219) = 3.96, $p < .001$. Students had a stronger understanding of the substance/property content. This provides one possible reason for why students who received faded written scaffolds during the unit had higher substance/property reasoning on the posttest, but not different chemical reaction reasoning than those who received continuous scaffolds. In order to effectively apply the scientific explanation model, they may have needed a sufficient understanding of the content.

Yet content knowledge alone does not appear to explain the different performances of the Faded and Continuous groups. We tested whether students' understanding of the content knowledge as measured by the multiple-choice items varied between the Faded and Continuous groups. We performed an ANCOVAs on students' posttest multiple-choice scores; Scaffold Treatment (continuous, faded) was the fixed factor and the pretest multiple-choice score was the covariate. The effect of the pretest covariate was significant. Yet the effect of Scaffold Treatment on students' posttest multiple-choice scores was not significant, F (1, 217) = 1.21, *ns*. This suggests that the higher reasoning scores of the Faded group were not simply the result of a stronger understanding of the content. Both an understanding of the content and scientific explanation appears to be important for student performance.

Student Difficulty with the Reasoning Component of Explanations

Another trend in our results was that the scaffolds had the strongest influence on the reasoning component of the explanations. We hypothesize that this greater influence is because students had the most difficulty with the reasoning component of explanation. Table 3 displays students claim, evidence, and reasoning scores for each of the treatments on the pretest and posttest. This table shows that claim and evidence are higher than reasoning, regardless of treatment. For example, on the posttest the average score across all the students for claim was 3.20, for evidence 2.73, and for reasoning 1.76 (Maximum score for all three components is 5.0). Students appear to have a difficult time articulating why their data counts as evidence to support their claim and backing up that link between the claim and evidence by including the appropriate scientific principles.

Discussion

A distinctive aim of science (Nagel, 1961) and science education (Driver et al., 2000; National Research Council, 2000) is to construct systematic and well-supported explanations. Our work suggests that providing students with an explicit instructional model for explanation and written explanation scaffolds results in students creating stronger explanations. Specifically, we were interested in whether the fading of the written supports resulted in greater student learning compared to students who received continuous written supports. Originally, Wood et al. (1976) described scaffolding as a flexible process that relies both on what a child knows and the learning task. Scaffolding should provide just enough information that learners can make progress independently (Hogan & Pressley, 1997), which suggests that scaffolds should be adjusted or faded as students learn (Stone, 1998). Although the fading of scaffolds has been found to be beneficial for students during teacher-student interactions (Palincsar & Brown, 1984), there has been little research examining whether written supports can be faded to promote student learning. We argue that in order for a support to be considered a scaffold, it ultimately needs to be able to fade. Our findings suggest that fading written prompts that include both context specific and generic explanation supports better equips students to write explanations when they are not provided with support. This suggests that even though written supports do not offer the individualization of one-on-one interactions, they can still be effective scaffolds in complex learning environments.

Fading written supports had a significant effect on students' test scores, but not on the explanations that they constructed during class. We conjecture that the written scaffolds did not influence student performance in class, because of the multiple other supports available to students within the complex learning environment. Students, tools and teachers work together as a system to support student learning (Reiser, 2004). Tabak (2004) argues that these different supports can be synergistic if they interact or work in concert to support students in a specific learning goal. Teacher practices are particularly important for the successful use of scaffolded cognitive tools (Pea, 2004). For example during the focal lesson of our study, students did not receive written scaffolds on their investigation sheets; rather, the teacher provided students with support. Consequently, the increases in student performance from pretest to the focal lesson (See Figures 2-4) were due to something other than the written scaffolds. We conjecture that this increase in student performance was due to teacher practices during the focal lesson such as making the framework of explanation explicit to students and modeling explanation construction, which supported students in the creation of their own explanations. In other research, we found that the instructional practices that a teacher engages in can significantly affect student learning of scientific explanations (Lizotte, McNeill & Krajcik, 2004; McNeill & Krajcik, in review). In considering the role of the written scaffolds, it is important to consider these other aspects of the classroom environment. In our future work, we hope to explore the interactions between teacher instructional practices and other scaffolded tools and how they influence student learning.

When students constructed explanations independently on the pre and posttest, the fading of scaffolds significantly effected students' reasoning scores. Students who received the faded scaffolds provided stronger reasoning in their scientific explanations where they showed why the data counted as evidence to support the claim by using appropriate scientific principles. We believe that the scaffolds had the strongest effect on students' reasoning scores because this was the most difficult component for students. Previous research has found that students have difficulty providing the backing for their claims and evidence both in their written explanations (Bell & Linn, 2000) and during classroom discussion (Jiménez-Aleixandre, et al., 2000). We found that students' reasoning scores were lower than both their claim, and evidence scores throughout the unit. We hypothesize that students' difficulty with reasoning made the scaffolds particularly important for this component of scientific explanation throughout the unit. By fading the scaffolds, this may have forced students to revisit this question of "What is the reasoning?" instead of relying on the support of the written scaffolds. Making the learning task more difficult in the short term by fading scaffolds encouraged greater student understanding and independence in the long term (Reiser, 2004). By fading the scaffolds we forced students to think about and apply what they had learned from the previous scaffolds to the current learning task. This may have better equipped them to provide the reasoning on the posttest when no scaffolds were provided and students had to complete the task independently.

We found a relationship between students' content knowledge and their ability to construct explanations. For the content areas where students' understanding was weaker, they also constructed weaker scientific explanations. Students' performance depends on both their reasoning capabilities and their understanding of the science content (Metz, 2000). Consequently, when students construct poor explanations it is difficult to tease out whether students' difficulty stems from their lack of understanding of the content or their lack of understanding of scientific explanations. In order to understand the nature or development of scientific reasoning, an individual's conceptual knowledge needs to be taken into consideration (Zimmerman, 2000). The relationship between content understanding and explanation construction suggests that one reason the scaffolds did not appear to influence students' explanations on the chemical reaction questions on the posttest may be because the students did not understand the chemical reaction content well enough to demonstrate their understanding of scientific explanations. Without a strong understanding of the chemical reaction content, they could not write a strong scientific explanation about chemical reactions. Yet content knowledge does not appear to explain the difference between the faded and continuous groups, as there was not a significant difference in their content knowledge as measured by their multiple-choice scores. This suggests that the faded groups higher reasoning scores for the substance and property explanations reflects a stronger ability to include reasoning in their scientific explanations.

Lee and Songer's (2004) study found that fifth and sixth grade students did not benefit from the fading of context specific supports. This suggests that fading alone does not determine the effectiveness of written instructional scaffolds. Other factors such as the age of the student, the content area, the role of the teacher as well as the type, language, and duration of the scaffolds influence when and how scaffolds should be faded. For example, we mentioned that we chose to fade the scaffolds using equal intervals, two written prompts in each stage. A different method of fading may in fact be more effective for student learning. Further research needs to be conducted to determine how various factors influence the effects of fading written scaffolds.

The format and language of the scaffolds may also affect student learning. Stone (1993) discusses how successful scaffolding involves the construction of shared definitions in a particular situation or context. Although Stone discusses the importance of a shared discourse in adult-child interactions, we believe that this shared definition is also important in written scaffolds. One of the reasons students had difficulty constructing the reasoning portion of their explanations may be because they did not share the same understanding of reasoning as we intended in the curriculum. In the written scaffolds, we described reasoning as "a statement that connects your evidence to your claim" (Figure 1). Students may not have understood what we meant by "connect". We wanted students to include the scientific principles that allowed them to use the evidence to support their claim, but the students may not have shared that same definition of reasoning. Using a different scaffolding format or language in the scaffold, may have been more effective. In our current work as well as that of our colleagues at Northwestern, we are experimenting with other variations of the wording of the explanation scaffolds. For example, we are currently investigating whether the context specific versus generic nature of the written prompts influences student learning. Our colleagues at Northwestern are exploring whether the wording of scaffolds as persuasive questions, instead of declarative statements, encourages students to defend their scientific explanations (Kuhn & Reiser, 2005).

Educators need to pay careful attention to the relationship between the type, language and duration of scaffolds for different content areas and inquiry practices. This relationship may depend on the difficulty of the science content, students' grade level or prior experiences, as well as other characteristics of the learning environment. Furthermore, scaffolding another inquiry practice, such as the design of experiments or asking questions, may have a different optimal scaffolding solution. Participating in other inquiry practices may present different difficulties for students, which would influence the design of the scaffolds.

The design of scaffolded instructional materials requires the careful consideration of multiple factors. Our work suggests that ultimately written supports should fade to scaffold students to obtain stronger understandings of inquiry practices that they can perform independently, such as scientific explanations. Fading scaffolds better prepares students to become scientifically literate in that they are able to engage in scientific inquiry practices such as the construction of scientific explanations. The goal of science education is to prepare students to participate in scientific practices outside of the science classroom (McGinn & Roth, 1999). When individuals are confronted with science in their everyday lives, there are no instructional supports to reduce the complexity of the task or provide the rationale of the inquiry practices. To prepare students for this lack of support in the everyday world instructional materials need to eventually fade scaffolds to better prepare students for legitimate participation.

Because scientific inquiry practices are so complex, they cannot be scaffolded all at once (Pea, 2004). Yet the design of inquiry curriculum should include a sequence of well-delineated inquiry practices that are part of the objectives of the curriculum (Kuhn, Black, Keselman & Kaplan, 2000). Our work suggests that fading scaffolds at the seventh grade level is appropriate for supporting students in the construction of scientific explanations consisting of claim, evidence, and reasoning. This is an important finding for a much larger learning progression where students will need to consider other aspects of explanation such as counter arguments and rebuttals, which may not be effectively faded at the middle school level. As a research community we need to provide more research around the scaffolding of explanation and other inquiry practices for the instructional design of a complete k-12 science curriculum that carefully sequences the development of scientific practices over time.

Acknowledgements

This research was conducted as part of the Investigating and Questioning our World through Science and Technology (IQWST) project and the Center for Curriculum Materials in Science (CCMS), supported in part by the National Science Foundation grants ESI 0101780 and ESI 0227557 respectively. Any opinions expressed in this work are those of the authors and do not necessarily represent either those of the funding agency or the University of Michigan. We would like to thank all of the researchers involved with IQWST and CCMS, especially Betsy Davis and Brian Reiser, for their feedback on this paper.

References

Aikenhead, G. S. (2004) Science-based occupations and the science curriculum: Concepts of evidence. *Science Education, 89 (2), 242-275*.

American Association for the Advancement of Science. (1993). *Benchmarks for science literacy.* New York: Oxford University Press.

Bell, P., & Linn, M. (2000). Scientific arguments as learning artifacts: Designing for learning from the web with KIE. *International Journal of Science Education. 22 (8),* 797-817.

Blumenfeld, P., Fishman, B. J., Krajcik, J., Marx, R. W., & Soloway, E. (2000). Creating usable 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.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, *18*(1), 32-41.

Burkhardt, H. & Schoenfeld, A. H. (2003). Improving educational research: Toward a more useful, more influential, and better-funded enterprise. *Educational Researcher, 32*(9), 3-14.

Chen, Z., & Klahr, D. (1999). All other things being equal: Acquisition and transfer of the control of variables strategy. *Child Development*. *70*(5). 1098-1120.

Chinn, C. A. & Brewer, W. F. (2001). Models of data: A theory of how people evaluate data. *Cognition and Instruction*, *19*, 323-393.

Chinn, C. & Brown, D. E. (2000). Learning in science: A comparison of deep and surface approaches. *Journal of Research in Science Teaching, 37*(2), 109-138.

Cho, K.-L., & Jonassen, D. H. (2002). The effects of argumentation scaffolds on argumentation and problem solving. *Educational Technology Research and Development*, 50(3), 5-22.

Cobb, P. (2002). Reasoning with tools and inscriptions. *The Journal of the Learning Sciences, 11*(2&3), 187-215.

Collin, A., Brown, J. S., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the craft of reading, writing, and mathematics. In L B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 453-494). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Davis, E. A. (2003). Prompting middle school science students for productive reflection: generic and directed prompts. *The Journal of the Learning Sciences. 12*(1), 91-142.

Davis, E. A., & Miyake, N. (2004). Explorations of scaffolding in complex systems. *The Journal of the Learning Sciences. 13*(3), 265-272.

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., Newton, P. & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education. 84(3),* 287-312.

Edelson, DC, Gordin, D.N., & Pea, R.D. (1999). Addressing the challenges of inquiry-based learning through technology and curriculum design*. Journal of the Learning Sciences,* 8(3/4) 391-450.

Erduran, S., Simon, S. & Osborne, J. (2004) TAPing into argumentation: Developments in the application of Toulmin's argument pattern for studying science discourse. *Science Education. 88*(6), 915-933.

Herrenkohl, L. R., Palincsar, A. S., DeWater, L. S., & Kawasaki, K. (1999). Developing scientific communities in classrooms: A Sociocognitive Approach. *The Journal of the Learning Sciences. 8*(3&4), 451-493.

Hogan, K. & Maglienti, M. (2001). Comparing the epistemological underpinnings of students and scientists' reasoning about conclusions. *Journal of Research in Science Teaching*. *38*(6). 663-687.

Hogan, K., & Pressley, M. (1997). *Scaffolding student learning: Instructional approaches and issues*. Cambridge, MA: Brookline Books.

Jiménez-Aleixandre, M. P., Rodríguez, A. B., & Duschl, R. A. (2000). "Doing the lesson" or "doing science": argument in high school genetics. *Science Education, 84*, 757-792.

Kolodner, J. L. & Nagel, K. (1999). The design discussion area: A collaborative learning tool in support of learning from problem-solving and design activities. In C. Hoadley & J. Roschelle (Eds.), *Proceedings of the conference of the computer supported collaborative learning* (pp. 300-307). Palo Alto, CA.

Koslowski, B. (1996). *Theory and evidence: The development of scientific reasoning.* Cambridge, MA: MIT Press.

Kuhn, D. (1991). *The skills of argument*. Cambridge, UK: Cambridge University Press.

Kuhn, D. (1993) Science as argument: Implications for teaching and learning scientific thinking. *Science Education, 77,* 319-338.

Kuhn, D., Black, J., Keselman, A., & Kaplan, D. (2000). The development of cognitive skills to support inquiry. *Cognition and Instruction, 18*(4), 495-523.

Kuhn, D., & Udell, W. (2001). The path to wisdom. *Educational Psychologist. 36 (4),* 261-264.

Kuhn, D & Udell, W. (2003). The development of argument skills. *Child Development. 74 (5),* 1245-1260.

Kuhn. L. & Reiser, B. (April, 2005). *Students constructing and defending evidence-based scientific explanations.* Paper presented at the annual meeting of the National Association for Research in Science Teaching, Dallas, TX.

Krajcik, J., Blumenfeld, P.C., Marx, R.W., Bass, K.M., & Fredericks, J. (1998). Inquiry in project based science classrooms: Initial attempts by middle school students. *Journal of the Learning Sciences*, 77, 317-337.

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.

Krajcik, J., Czerniak, C. M., & Berger, C. (1999). *Teaching children science: A project-based approach*. Boston, MA: McGraw-Hill.

Lee, H.-S. & Songer, N. B. (2004, April). *Longitudinal knowledge development: Scaffolds for Inquiry.* Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Lee, O. & Fradd, S. H. (1998). Science for all, including students from non english-language backgrounds. *Educational Researcher*, 27 (4), 12-21.

Lizotte, D. J, Harris, C. J., McNeill, K. L., Marx, R. W., & Krajcik, J. (2003, April). *Usable assessments aligned with curriculum materials: Measuring explanation as a scientific way of knowing.* Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Lizotte, D. J., McNeill, K. L., & Krajcik, J. (2004). Teacher practices that support students' construction of scientific explanations in middle school classrooms. In Y. Kafai, W. Sandoval, N. Enyedy, A. Nixon & F. Herrera (eds.), *Proceedings of the sixth international conference of the learning sciences* (pp. 310-317). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Metz, K. E. (2000). Young children's inquiry in biology: Building the knowledge bases to empower independent inquiry. In J. Minstrell & E. H. van Zee (eds.), *Inquiry into inquiry learning and teaching in science* (pp. 371-404). Washington, DC: American Association for the Advancement of Science.

McGinn, M. K. & Roth, W-M. (1999). Preparing students for competent scientific practice: Implications of recent research in science and technology studies. *Educational Researcher, 28*(3) 14-24.

McNeill, K. L., Harris, C. J., Heitzman, M., Lizotte, D. J., Sutherland, L. M., & Krajcik, J. (2004). How can I make new stuff from old stuff? In J. Krajcik & B. J. Reiser (Eds.), *IQWST: Investigating and questioning our world through science and technology*. Ann Arbor, MI: University of Michigan.

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. & Krajcik, J. (in review). Scientific explanations: Characterizing and evaluating the effects of teachers' instructional practices on student learning.

McNeill, K. L., Lizotte, D. J, Harris, C. J., Scott, L. A., Krajcik, J., & Marx, R. W. (2003, March). *Using backward design to create standards-based middle-school inquiry-oriented chemistry curriculum and assessment materials*. Paper presented at the annual meeting of the National Association for Research in Science Teaching, Philadelphia, PA.

Nagel, E. (1961). *The structure of science: Problems in the logic of science education.* New York, NY: Harcourt, Brace, & World, Inc.

National Research Council. (1996). *National science education standards.* Washington, DC: National Academy Press.

National Research Council. (2000). *Inquiry and the national science education standards: A guide for teaching and learning.* Washington, D.C.: National Academy Press.

Newton, P., Driver, R., & Osborne, J. (1999). The place of argumentation in the pedagogy of school science. *International Journal of Science Education*, 21(5), 553-576.

Osborne, J. Collins, S., Ratcliffe, M., Millar, R. & Duschl, R. (2003). What "Ideas-about-science" should be taught in school science? A Delphi study of the expert community. *Journal of Research in Science Teaching*, *40*(7), 692-720.

Osborne, J., Erduran, S. & Simon, S. (2004). Enhancing the quality of argumentation in school science. *Journal of Research in Science Teaching, 41*(10), 994-1020.

Owensby, J. N. & Kolodner, J. L. (2004). Case application suite: Scaffolding use of expert cases in middle-school project-based inquiry classrooms. In Y. Kafai, W. Sandoval, N. Enyedy, A. Nixon & F. Herrera (eds.), *Proceedings of the sixth international conference of the learning sciences* (pp. 310-317). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Palincsar, A. S. (1998). Keeping the metaphor of scaffolding fresh – A response to C. Addison Stone's "The metaphor of scaffolding: Its utility for the field of learning disabilities." *Journal of Learning Disabilities, 31*(4), 370-373.

Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction*, 1, 117-175.

Passmore, C. & Stewart, J. (2002). A modeling approach to teaching evolutionary biology in high schools. *Journal of Research in Science Teaching, 39*(3), 185-204.

Pea, R. D. (2004). The social and technological dimensions of scaffolding and related theoretical concepts for learning, education, and human activity. *The Journal of the Learning Sciences. 13*(3), 423-451.

Perkins, D. N. & Salomon, G. (1989). Are cognitive skills context-bound? *Educational Researcher*, *18*(1), 16-25.

Quintana, C., Reiser, B., Davis, E. A., Krajcik, J., Fretz, E., Golan, R, Kyza, E., Edelson, D. & Soloway, E.. (2004). A scaffolding design framework for designing educational software. *The Journal of the Learning Sciences*, *13*(3), 337-386.

Reiser, B. J. (2004). Scaffolding complex learning: The mechanisms of structuring and problematizing student work. *The Journal of the Learning Sciences. 13*(3), 273-304.

Reiser, B. J., Krajcik, J., Moje, E. B., & Marx, R. W. (2003, March). *Design strategies for developing science instructional materials*. Paper presented at the Annual Meeting of the National Association for Research in Science Teaching, Philadelphia, PA.

Reiser, B., Tabak, I., Sandoval, W., Smith, B., Steinmuller, F., & Leone, A. (2001). BGuILE: Strategic and conceptual scaffolds for scientific inquiry in biology classrooms. In S.M. Carver & D. Klahr (Eds.), *Cognition and instruction: Twenty-five years of progress* (pp. 263-305). Mahwah, NJ: Erlbaum.

Reznitskaya, A., & Anderson, R. C. (2002). The argument schema and learning to reason. In C. C. Block, & M. Pressley (Eds.), *Comprehension instruction: Research-based best practices* (pp. 319-334). New York: The Guilford Press.

Reznitskaya, A., Anderson, R. C., McNurlen, B., Nguyen-Jahiel, K., Archodidou, A., & Kim, S.-Y. (2001). Influence of oral discussion on written argument. *Discourse Processes*, *32*(2&3), 155-175.

Sadler, T. D. (2004). Informal reasoning regarding socioscientific issues: A critical review of research. *Journal of Research in Science Teaching*. *41*(5). 513-536.

Sandoval, W. (2003). Conceptual and epistemic aspects of students' scientific explanations. *The Journal of the Learning Sciences, 12*(1), 5-51.

Sandoval, W. A. & Millwood, K. A. (2005). The quality of students' use of evidence in written scientific explanations. *Cognition and Instruction, 23*(1), 23-55.

Sandoval, W. A. & Reiser, B. J. (2004). Explanation-driven inquiry: Integrating conceptual and epistemic scaffolds for scientific inquiry. *Science Education, 88*, 345-372.

Scardamalia, M. and Bereiter, C. (1991). Higher levels of agency for children in knowledge building: A challenge for the design of new knowledge media. *The Journal of the Learning Sciences 1*(1), 37-68.

Stone, C. A. (1993). What is missing in the metaphor of scaffolding? In E. A. Forman & N. Minick & C. A. Stone (Eds.), *Contexts for learning: Sociocultural dynamics in children's development* (pp. 169-183). New York: Oxford University Press.

Stone, C. A. (1998). The metaphor of scaffolding: its utility for the field of learning disabilities. *Journal of Learning Disabilities, 31*(4), 344-364.

Tabak, I. (2004). Synergy: A complement to emerging patterns in distributed scaffolding. *Journal of the Learning Sciences*. *13*(3), 305-335.

Toulmin, S. (1958). *The uses of argument.* Cambridge, UK: Cambridge University Press.

Van Eemeren, F. H., Grootendorst, R., Henkemans, F. S., Blair, J. A., Johnson, R. H., Krabbe, E. C. W., Plantin, C., Walton, D. N., Willard, C. A., Woods, J., & Zarefsky, D. (1996). *Fundamentals of argumentation theory: A handbook of historical backgrounds and contemporary developments*. Mahwah, NJ: Lawrence Erlbaum Associates.

Voss, J. F. & Van Dyke, J. A. (2001). Argumentation in Psychology: Background Comments. *Discourse Processes*, 32(2&3), 89-111.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. London, England: Cambridge University Press.

White, B., & Frederiksen, J. (1998). Inquiry, modeling, and metacognition: Making science accessible to all students. *Cognition and Instruction, 16*(1), 3-118.

White, B., & Frederiksen, J. (2000). Metacognitive facilitation: An approach to making scientific inquiry accessible to all. In J. Minstrell & E. v. Zee (Eds.), *Inquiring into Inquiry Learning and Teaching in Science* (pp. 283-315). Washington D.C.: AAAS.

Wood, D., Bruner, J. S. & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry. 17,* 89-100.

Yeh, S. S. (2001). Tests worth teaching to: Constructing state-mandated tests that emphasize critical thinking. *Educational Researcher, 30*(9), 12-17.

Zembal-Saul, C., Munford, D., Crawford, B., Friedrichsen, P. & Land, S. (2002). Scaffolding preservice science teachers' evidence-based arguments during an investigation of natural selection. *Research in Science Education*, *32(4),* 437-465.

Zimmerman, C. (2000). The development of scientific reasoning skills. *Developmental Review, 20*, 99-149.

Zohar, A., & Nemet, F. (2002). Fostering students' knowledge and argumentation skills through dilemmas in human genetics. *Journal of Research in Science Teaching*. *39*(1), 35-62.

Appendix A: Base Rubric

Base Explanation Rubric

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

Appendix B: Specific Explanation Rubric for Test Question #1

Component		Level			
		0	**1a**	**1b**	**2**
Claim	**Criteria:**	Does not make a claim, or makes an inaccurate claim.	Makes an accurate but incomplete claim that some of the solids are the same substance.	[Does not apply.]	Makes an accurate and complete claim that solids 2 and 4 are the same substance.
	Student Exemplar(s):	"Solids 1 and 4 are the same."	"Some of the solids are the same."		"There are two solids that are the same. Them two solids are 2 and 4."
Evidence	**Criteria:**	Does not provide evidence, or only provides inappropriate evidence such as mass.	1 piece of appropriate evidence (i.e. solubility, melting point or color). May include some inappropriate evidence (i.e. mass).	2 pieces of appropriate evidence (i.e. solubility, melting point or color). May include some inappropriate evidence (i.e. mass). OR: 3 pieces of appropriate evidence and inappropriate evidence.	3 pieces of appropriate evidence (i.e. solubility, melting point and color) and no inappropriate evidence
	Student Exemplar(s):	"They have the same mass."	"Solid 2 and 4 same soluble in water. Solid 2 and 4 have the same mass."	"Both 2 and 4 are white and have a melting point of 175 °C."	2 and 4 have the same solubility, melting point is 175 °C and they are white."
Reasoning	**Criteria:**	Does not provide reasoning, or only provides reasoning that does not link evidence to claim.	Provides reasoning that links the claim and evidence by repeating the evidence.	Provides reasoning that includes an insufficient generalization about properties (i.e. states that density is a property).	Provides reasoning that includes an appropriate and sufficient generalization about different substances having different properties.
	Student Exemplar(s):	"My reasoning that supports my claim about whether my answer is right or not is the chart because it shows all my evidence."	"Since they have the same solubility, melting point, and color, they are the same substance." [repeats evidence]	"Mass is not a property so that don't matter. Only soluble in water and melting point and color matter. These are some of the most important properties so they matter the most."	"… soluble in water, melting point and color are all properties of a substance and solid 2 and solid 4 have the same properties so they are the same substance."

Appendix C: Scientific Explanation Test Questions

Question #1: Examine the following data table:

	Mass	**Soluble in Water**	**Melting Point**	**Color**
Solid 1	65 g	Yes	136 °C	yellow
Solid 2	38 g	Yes	175 °C	white
Solid 3	21 g	No	89 °C	white
Solid 4	65 g	Yes	175 °C	white

Write a **scientific explanation** that answers the question: Are any of the solids the same substance?

Question #4: Carlos has two liquids, butanic acid and butanol. He determines a number of measurements for the two liquids and then mixes them together. After heating and stirring the liquids, they form two separate layers, layer A and layer B. Carlos uses an eyedropper to take a sample from each layer, and he determines a number of measurements for each.

	Volume	**Mass**	**Density**	**Solubility in water**	**Melting Point**
Butanic acid	10.18 cm^3	9.78 g	0.96 g/cm^3	Yes	-7.9 °C
Butanol	10.15 cm^3	8.22 g	0.81 g/cm^3	Yes	-89.5 °C
Layer A	2 cm^3	1.74 g	0.87 g/cm^3	No	-91.5 °C
Layer B	2 cm^3	2.0 g	1.0 g/cm^3	Yes	0 °C

Write a **scientific explanation** that answers the question: Did a chemical reaction occur when Carlos mixed butanic acid and butanol?

Figure 1. Example Explanation Item Showing Generic and Context-Specific Portions

---

Item:
Write a scientific explanation that answers the question: Are the nail and the wrench the same substance or different substances?

Claim (Write a sentence that states whether the nail and wrench are the same or different substances.)

Two Pieces of Evidence (Provide two pieces of data that support your claim that the nail and the wrench are the same or different substances.)

Evidence #1

Evidence #2

Reasoning (Write a statement that connects your evidence to your claim that the nail and the wrench are the same or different substances.)

Note: Circled text indicates the generic portion of the scaffolds; non-circled text indicates the context-specific portion of the scaffolds.

Table 1: Continuous and Faded Scaffolding Treatments Over the Three Stages

Stage	Continuous Scaffold	Faded Scaffold
Focal Lesson	No scaffolds	No scaffolds
Stage I	**Claim** (Write a sentence that states whether the nail and wrench are the same or different substances.) **Two Pieces of Evidence** (Provide two pieces of data that support your claim that the nail and the wrench are the same or different substances.) **Evidence #1** **Evidence #2** **Reasoning** (Write a statement that connects your evidence to your claim that the nail and the wrench are the same or different substances.)	**Claim** (Write a sentence that states whether the nail and wrench are the same or different substances.) **Two Pieces of Evidence** (Provide two pieces of data that support your claim that the nail and the wrench are the same or different substances.) **Evidence #1** **Evidence #2** **Reasoning** (Write a statement that connects your evidence to your claim that the nail and the wrench are the same or different substances.)
Stage II	**Claim** (Write a sentence that states whether or not boiling is a chemical reaction.) **Two Pieces of Evidence** (Provide two pieces of data that support your claim whether or not boiling is a chemical reaction.) **Evidence #1** **Evidence #2** **Reasoning** (Write a statement that connects your evidence to your claim whether or not boiling is a chemical reaction.)	**Claim** (Write a sentence that answers the question.) **Evidence** (Provide data that support your claim.) **Reasoning** (Connect evidence to claim.)
Stage III	**Claim** (Write a sentence that states whether mass stayed the same or changed.) **One Piece of Evidence** (Provide one piece of data that supports your claim whether mass stayed the same or changed.) **Evidence #1** **Reasoning** (Write a statement that connects your evidence to your claim whether mass stayed the same or changed.)	Remember to include claim, evidence, and reasoning.

Table 2: Timeline and Description of Explanations During the Stuff Unit

Stage	Day*	Title of Learning Task	Description of Learning Task
Focal Lesson	12	Activity 5.1: Are fat and soap the same or different substances?	Students collect data from their previous investigations where they determined different properties of fat and soap. They construct explanations on whether soap and fat are the same or different substance.
Stage I	13	Reader: Are these two objects the same substance or different substances?	Students are provided with data for the properties of a nail and a wrench. They construct explanations on whether they are the same substance.
	15	Activity 7.1: What happens to properties when I combine stuff?	Students investigate what happens when they combine substances. They construct explanations on whether a new substance is formed.
Stage II	18	Reader: A closer look at electrolysis	In class, students investigate electrolysis where they separate water into hydrogen and oxygen. Students are provided with the properties for the substances and have to construct explanations on how they know a chemical reaction occurred.
	20	Activity 9.1: Does boiling water make new stuff?	Students investigate and construct explanations on whether boiling is a chemical reaction.
Stage III	29	Activity 13.1: Does mass change when Alka-Seltzer reacts?	Students react Alka Seltzer and water in an open container and determine whether the mass changes. They construct explanations on whether mass changes in a chemical reaction.
	30	Activity 13.1: Does mass really change when Alka-Seltzer reacts?	Students react Alka Seltzer and water in a closed container and determine whether the mass changes. They construct explanations on whether mass changes in a chemical reaction.

* The instructional materials suggest that the unit should take a total of 36 class periods. All teachers in this study had regular class periods (not block) so one class period equaled one day. This is the suggested day during the unit the explanation would be completed, though there was some variation across teachers' enactments.

Table 3: Faded and continuous performance for claim, evidence, and reasoning

Treatment	Pretest *M* (*SD*)[a]	Posttest *M* (*SD*)	*t*-Value[b]	Effect Size[c]
Faded (n = 123)				
Claim	2.37 (1.46)	3.23 (1.61)	5.77***	0.59
Evidence	1.47 (1.21)	2.47 (1.59)	7.47***	0.83
Reasoning	0.39 (0.63)	1.92 (1.62)	11.12***	2.43
Continuous (n = 97)				
Claim	1.88 (1.35)	3.17 (1.46)	9.36***	0.96
Evidence	1.10 (1.06)	2.25 (1.49)	7.97***	1.08
Reasoning	0.42 (0.66)	1.55 (1.45)	7.97***	1.71

[a] Maximum Score = 5.0
[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.
*** $p < .001$

Figure 2. Claim Scores Over the Unit

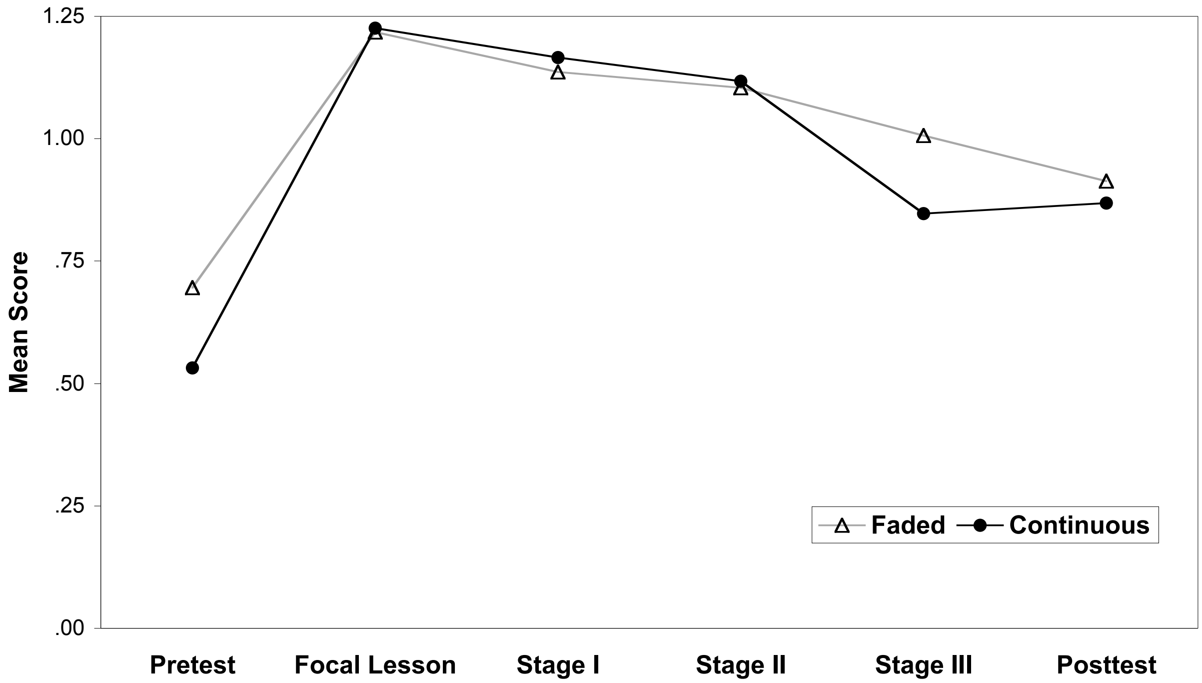

Figure 3. Evidence Scores Over the Unit

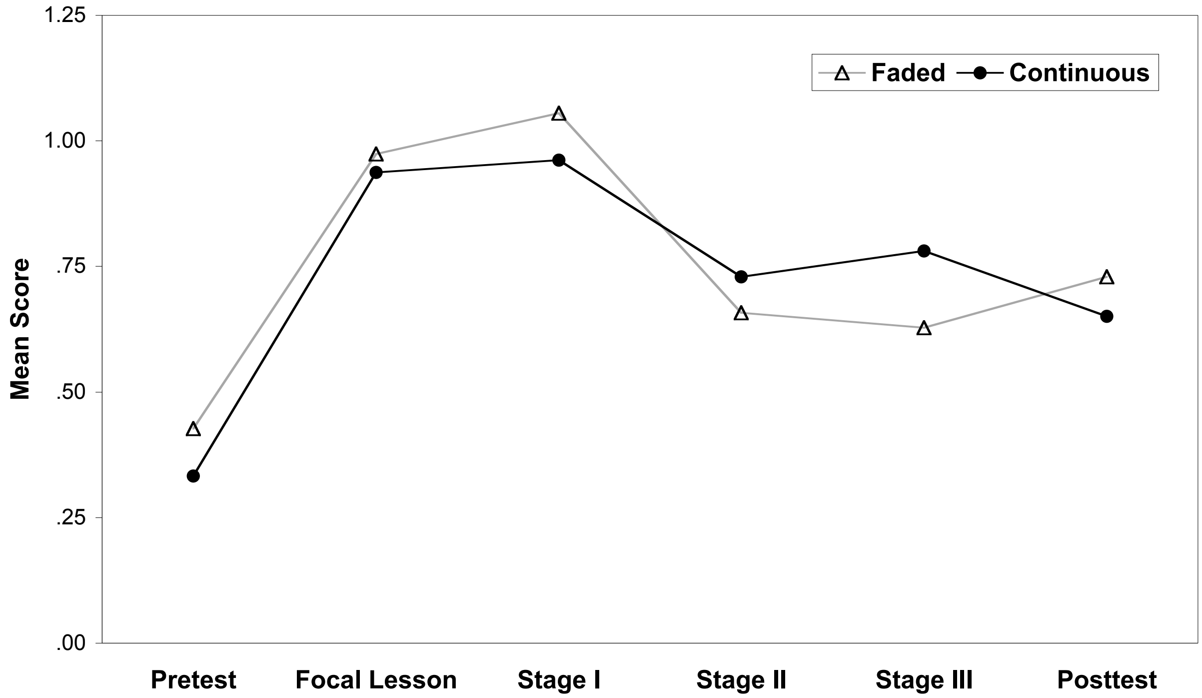

Figure 4. Reasoning Scores Over the Unit

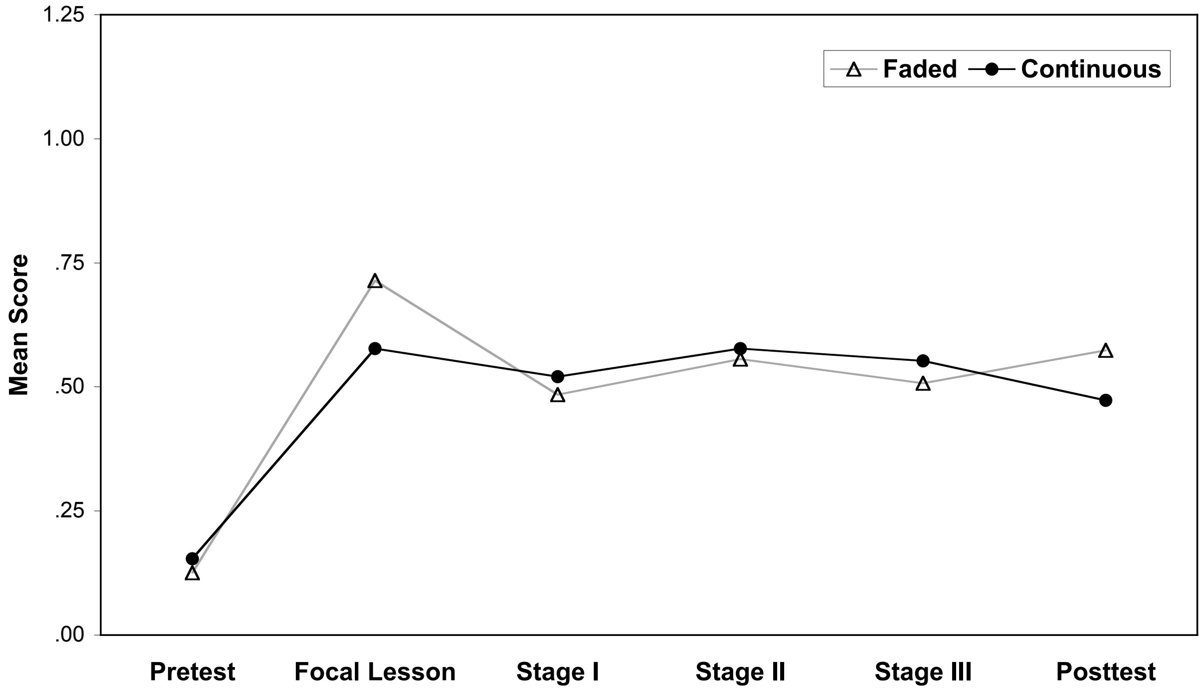

Figure 5 Influence of Scaffold Treatments on Posttest Reasoning for All Content

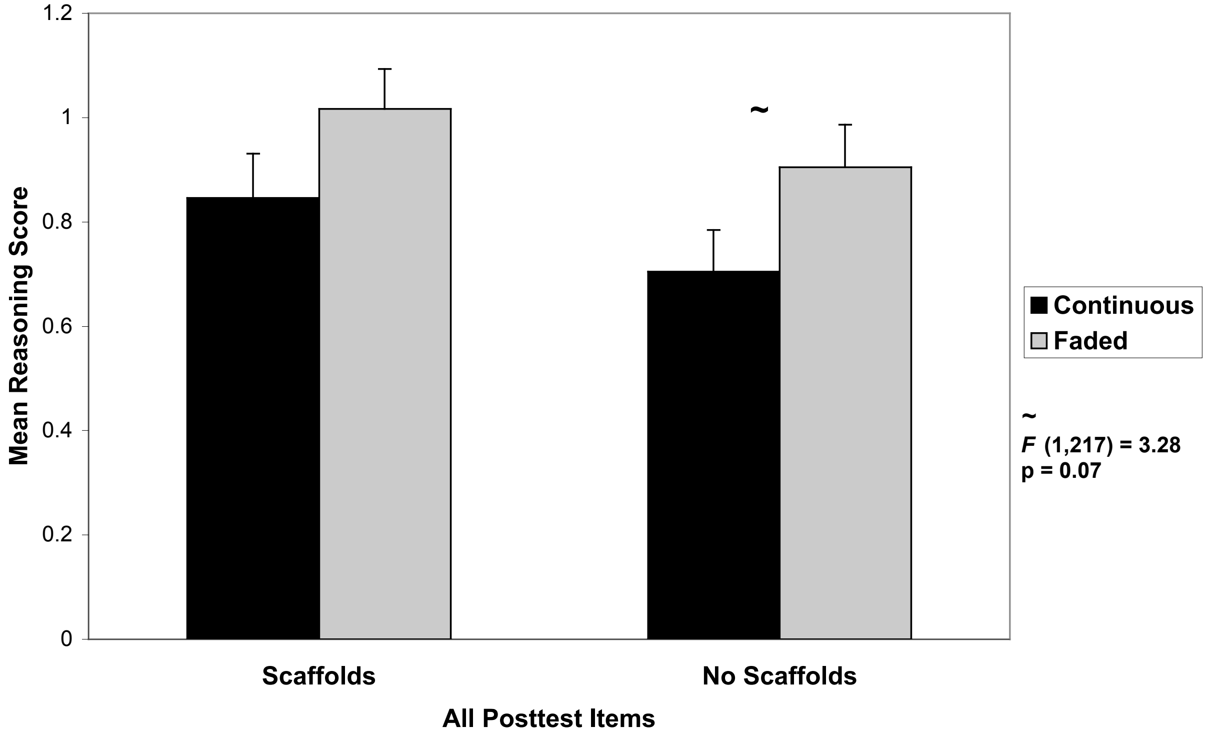

Figure 6. Influence of Scaffold Treatments on Posttest Reasoning for Substance/Property Items

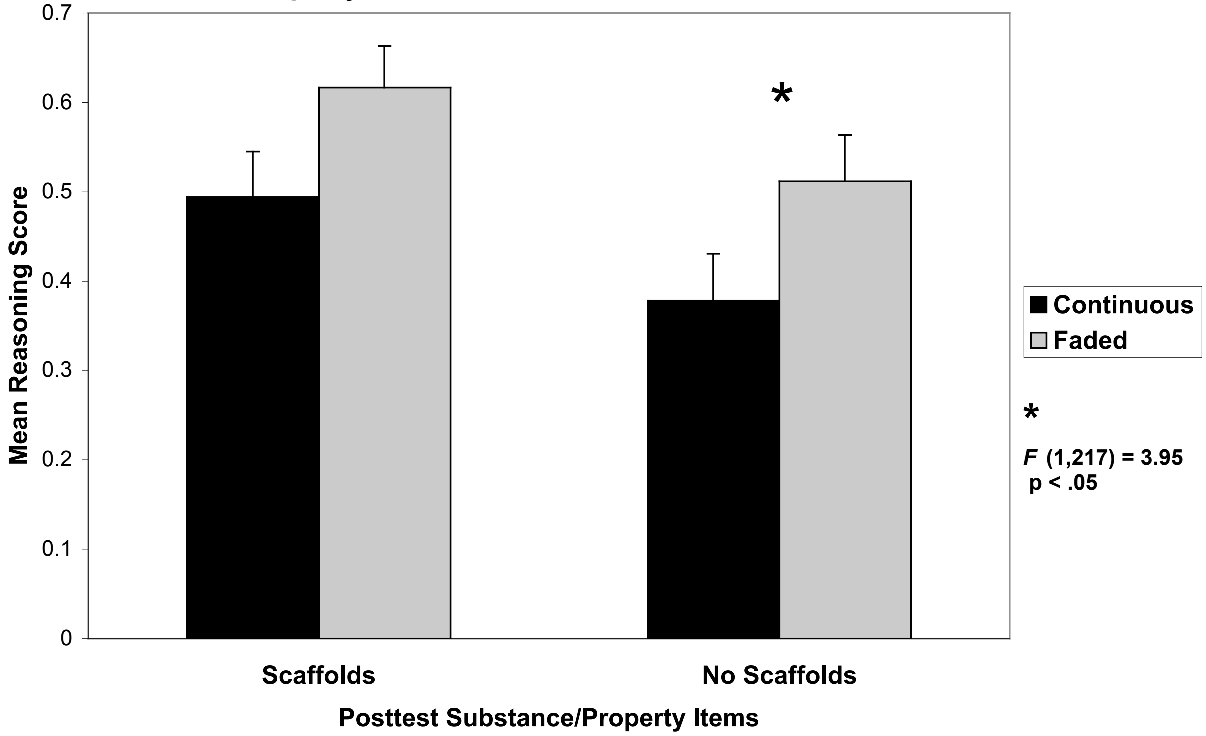

Figure 7
Influence of Test Item Scaffolds on Reasoning of Students in the Faded Group

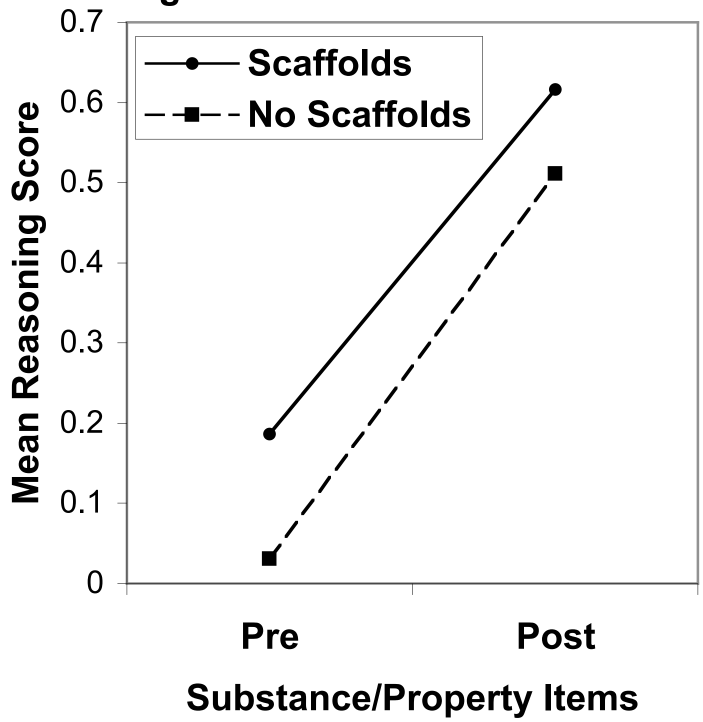

Figure 8
Influence of Test Item Scaffolds on Reasoning of Students in the Continuous Group

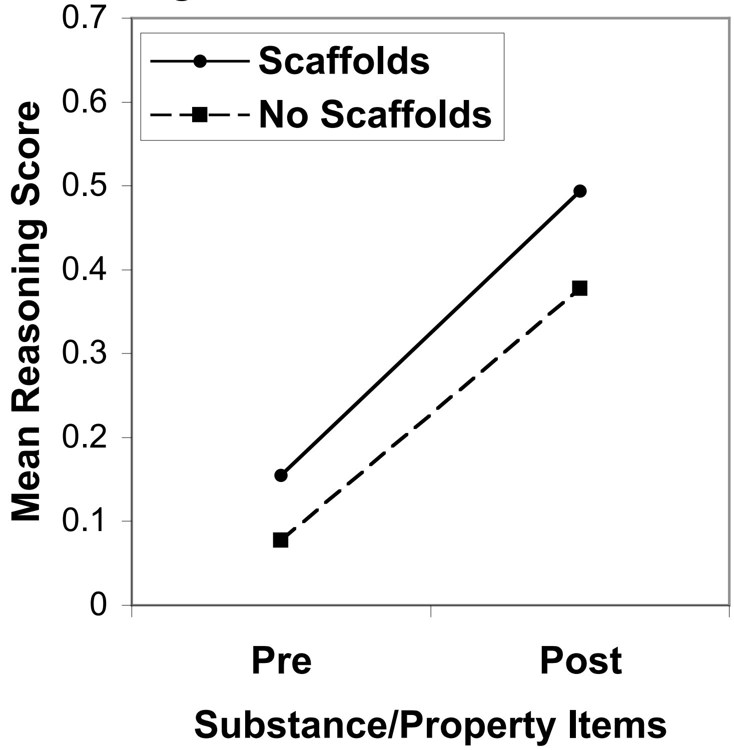

Figure Captions

Technical Notes

---

[1] After scoring the explanations, we converted the rubric levels to numeric scores for statistical analysis (e.g. level 0 = 0, level 1a = 5/12, level 1b = 10/12, and level 2 = 15/12).

[2] For all ANCOVAs presented, the effects of covariates are significant unless otherwise noted. Statistics for significant covariates are omitted.

[3] Both substance/property items and chemical reaction items were equally weighted for a maximum possible score of 5.0.