JOURNAL OF RESEARCH IN SCIENCE TEACHING VOL. 45, NO. 1, PP. 53–78 (2008)

Scientific Explanations: Characterizing and Evaluating the Effects of Teachers' Instructional Practices on Student Learning

Katherine L. McNeill,[1] Joseph Krajcik[2]

[1]*Lynch School of Education, Boston College, 140 Commonwealth Avenue, Chestnut Hill, Massachusetts 02467*

[2]*School of Education, University of Michigan, Ann Arbor, Michigan*

Received 13 June 2006; Accepted 16 January 2007

Abstract: Teacher practices are essential for supporting students in scientific inquiry practices, such as the construction of scientific explanations. In this study, we examine what instructional practices teachers engage in when they introduce scientific explanation and whether these practices influence students' ability to construct scientific explanations during a middle school chemistry unit. Thirteen teachers enacted a project-based chemistry unit, *How can I make new stuff from old stuff?*, with 1197 seventh grade students. We videotaped each teacher's enactment of the focal lesson on scientific explanation and then coded the videotape for four different instructional practices: modeling scientific explanation, making the rationale of scientific explanation explicit, defining scientific explanation, and connecting scientific explanation to everyday explanation. Our results suggest that when teachers introduce scientific explanation, they vary in the practices they engage in as well as the quality of their use of these practices. We also found that teachers' use of instructional practices can influence student learning of scientific explanation and that the effect of these instructional practices depends on the context in terms of what other instructional practices the teacher uses. © 2007 Wiley Periodicals, Inc. J Res Sci Teach 45: 53–78, 2008
Keywords: chemistry; inquiry; middle school science; classroom research

Classrooms are complex systems where many factors influence student learning, including teachers, peers, and other resources (Lampert, 2002). Recent research (Reiser et al., 2001) and reform documents (American Association for the Advancement of Science, 1993; National Research Council, 1996) argue that the role of the teacher is essential in structuring and guiding students' understanding of scientific inquiry, a key learning goal in recent science education reform efforts. Teachers need to support students in making sense of these scientific inquiry practices (Driver, Asoko, Leach, Mortimer, & Scott, 1994). We are interested in how different teacher instructional practices during the enactment of the same instructional unit influence

Contract grant sponsor: National Science Foundation; Contract grant number: ESI 0101780, ESI 0227557.
Correspondence to: K.L. McNeill; E-mail: kmcneill@bc.edu
DOI 10.1002/tea.20201
Published online 31 July 2007 in Wiley InterScience (www.interscience.wiley.com).

students' ability to engage in one important scientific inquiry practice, the construction of scientific explanations.

Role of Teachers in Inquiry

It is not enough to acknowledge that teachers play a critical role. We need to know what their role is in order to help support them in the difficult task of creating an inquiry-oriented classroom. Teachers have difficulty helping students with scientific inquiry practices, such as asking thoughtful questions, designing experiments, and drawing conclusions from data (Marx, Blumenfeld, Krajcik, & Soloway, 1997). Many science teachers may not have the appropriate expertise to create an inquiry-based learning environment (Krajcik, Mamlok, & Hug, 2001). Teachers need to learn new ways of teaching to promote scientific inquiry, which may differ from their own earlier socialization into school science as students (Lee, 2004; Metz, 2000). Although teachers often have difficulty supporting students, there is little research that provides guidance on what types of teacher practices may help students with scientific inquiry.

Research literature about inquiry classrooms often does not describe the classroom practices, rather classroom inquiry is summarized as "doing science," "hands-on science," or "real-world science" (Crawford, 2000). Furthermore, researchers often label a classroom as inquiry-oriented based on the nature of the curriculum materials used by the teacher and not by what the teacher and students are actually doing (Flick, 1995). Because teachers' beliefs about the nature of science, student learning, and the role of the teacher substantially affect their enactment of inquiry curriculum (Keys & Bryan, 2001), this raises the question of how using inquiry materials actually translates into inquiry-oriented classrooms. There is probably a range of inquiry occurring in these research studies labeled as exploring inquiry-oriented classrooms. Like other researchers (Flick, 2000; Keys & Bryan, 2001), we argue that there are few research studies that actually examine teachers' instructional practices in inquiry classrooms.

Scientific Explanations

One prominent scientific inquiry practice in both the standards documents (AAAS, 1993; NRC, 1996) and recent research literature in science education is the construction of scientific explanations or arguments (e.g., Bell & Linn, 2000; Driver, Newton, & Osborne, 2000; Jiménez-Aleixandre, Rodríguez, & Duschl, 2000; Kelly & Takao, 2002; Sandoval, 2003; Zohar & Nemet, 2002). Explanations refer to how or why a phenomenon occurs (Chin & Brown, 2000). An argument is an assertion with a justification (Kuhn, 1991) or a standpoint that is justified or defended for a particular audience (Van Eemeren et al., 1996). In our work, we use the word "explanation" to align with the national and state science standards that our teachers need to address, but our work builds on literature for both explanation and argumentation. Our goal is to help students construct scientific explanations about phenomena where they justify their claims using appropriate evidence and scientific principles.

Engaging students in scientific explanation and argumentation is a fundamental aspect of scientific inquiry (Duschl & Osborne, 2002). 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). Helping students engage in this practice may help shift their view of science away from science as a static set of facts to science as a social process where knowledge is constructed. Bell and Linn (2000) found that there is a correlation between students' views about science and the arguments that they construct. They suggested that engaging students in this practice may help refine their image of science. Furthermore, engaging in scientific explanation may help students construct a deeper understanding of the content knowledge. For example, Zohar and Nemet (2002) found that students who were engaged in a unit on argumentation skills through dilemmas in human genetics learned greater biological content knowledge than a comparison group who learned genetics in a more traditional manner.

Although engaging in scientific explanation is an important learning goal for students, students often have difficulty articulating and defending their knowledge claims (Sadler, 2004). Kuhn (1991) investigated both children and adults' ability to construct arguments and found that this practice often did not come naturally to them. They often had difficulty coordinating their claims and evidence. Even in a classroom setting where scientific explanation is an explicit goal, students still have many difficulties. Students can have difficulty using appropriate evidence (Sandoval, 2003) and providing sufficient evidence for their claims (Sandoval & Millwood, 2005). Students also have difficulty justifying why they chose their evidence to support their claims (Bell & Linn, 2000). In our previous work, we found that students had the most difficulty using scientific principles to justify why their evidence supports their claim (McNeill et al., 2006).

To help middle school students and teachers with this difficult scientific inquiry practice, we developed an instructional model for scientific explanation by adapting Toulmin's (1958) model of argumentation. The scientific explanation framework includes three components: a claim (a conclusion about a problem); evidence (data that supports the claim); and reasoning (a justification, built from scientific principles, for why the evidence supports the claim). In other work, we discussed the development of our framework as an instructional model (McNeill, Lizotte, Krajcik, & Marx, 2006; Moje et al., 2004) and as an assessment tool (McNeill & Krajcik, 2007). In this study, we explore how teachers' different uses of the explanation framework in their classrooms influenced student learning.

Teacher Instructional Practices Supporting Scientific Explanation

Few research studies have explored teacher instructional practices and their influence on students' construction of scientific explanation or argument. Previous research on students' construction of explanations in science has focused on scaffolds provided in the student materials or software programs (e.g., Bell & Linn, 2000; Lee & Songer, 2004; Sandoval, 2003; Zembal-Saul, Munford, Crawford, Friedrichsen, & Land, 2002) or on students' discussions in order to characterize their explanations (Jiménez-Aleixandre et al., 2000; Meyer & Woodruff, 1997).

Tabak (2004) looked at the role of the teacher in helping students construct evidence-based explanations. She argued that the teacher plays an important role in distributed scaffolding where many aspects of the learning environment, including software and other tools, come together synergistically to support student learning. Osborne, Erduran, & Simon, (2004) recently began exploring pedagogical practices that support students in argumentation. They argued that argumentation does not come naturally to students and that pedagogical practices are important for enhancing the quality of students' arguments. One of their initial findings is that teacher differences in their emphasis on components of argument may be a result of their different understandings of what counts as an argument (Erduran, Simon, & Osborne, 2004).

To further understand the role of teachers in supporting scientific explanation, we examined the literature for instructional practices that may support student learning of scientific explanation, but also other scientific inquiry practices, such as asking questions and designing experiments. From this literature, as well as a preliminary study we conducted on teacher practices (Lizotte, McNeill, & Krajcik, 2004), we decided to examine how teachers used four instructional practices during their introduction of scientific explanation: defining scientific explanation, making the rationale of scientific explanation explicit, modeling scientific explanation, and connecting scientific explanation to everyday explanation. We describe each of these instructional practices and provide examples of how they may support students' successful engagement in scientific explanations.

Defining Scientific Explanation

What is meant by various inquiry practices, such as designing experiments, asking questions, or constructing explanations, is not necessarily understood by students. One instructional practice a teacher may use to help students with these inquiry practices is to explicitly make the definition of these practices clear to students. Making scientific thinking strategies explicit to students can help facilitate their understanding and use of the strategies (Herrenkohl, Palincsar, DeWater, & Kawasaki, 1999). For example, Metz (2000) found that being explicit about scientific inquiry practices was important for helping children with the inquiry practice of formulating and refining questions. Explicit instruction may benefit diverse learners who are more likely to be unfamiliar with the participation rules and practices that are an essential part of scientific inquiry (Fradd & Lee, 1999). Consequently, this type of explicitness may allow students with impoverished experiences in science education to more effectively participate in classroom instruction as well as be beneficial to all students.

In terms of scientific explanation, students may create stronger explanations if teachers explicitly define what is meant by a scientific explanation and define the three components, claim, evidence, and reasoning. In a preliminary study (Lizotte et al., 2004), we found that when teachers explicitly defined scientific explanation, particularly the reasoning component, their students constructed stronger explanations.

Making the Rationale of Scientific Explanation Explicit

Instruction should both facilitate students' ability to perform inquiry practices and their understanding of the logic behind the practice (Kuhn, Black, Keselman, & Kaplan, 2000). Helping students understand the rationale behind why a particular scientific inquiry practice is important in science may result in students being better able to complete a performance. Chen and Klahr (1999) found that providing students with the rationale behind controlling variables in science experiments resulted in greater learning of this inquiry practice relative to students who did not receive the explicit instruction. Discussing why it is important to control variables to conduct a ‘‘fair’’ experiment helped students when they had to conduct their own experiments.

For scientific explanations, it may help students to construct stronger explanations if they understand why an individual may want to construct a scientific explanation and why providing evidence and reasoning results in a stronger, more convincing explanation. Students may need help understanding why someone would argue for a claim. Furthermore, it might be unclear why providing evidence and reasoning provides greater support than just providing an opinion.

Modeling Scientific Explanation

Modeling various inquiry practices is another instructional practice teachers can use to support student inquiry. Crawford (2000) argued that one of the key characteristics of a teacher establishing an inquiry-based learning environment is modeling the behaviors of a scientist. For example, the teacher Crawford researched in her case study frequently modeled how to grapple with data—specifically, through the extensive questioning of both the methods and results of data collection. Tabak and Reiser (1997) also found that student learning through collaboration in inquiry settings is more effective when teachers model strategies. For example, a teacher modeling how to reason from biological data can help students complete this same process of analyzing data on their own (Tabak, 2004).

Modeling how to include evidence and reasons for claims can help students in their own practice (Crawford, Kelly, & Brown, 2000). This can also help students learn how to use the general scientific explanation framework in a domain-specific context. Teachers can model explanations either through writing or speaking to provide students with concrete examples. Providing students with examples of strong and weak arguments can help them develop an understanding of what counts as a good argument (Osborne et al., 2004).

Connecting Scientific Explanation to Everyday Explanation

Connecting scientific discourse and inquiry practices to students' everyday discourse can help support students' learning of scientific inquiry. Lee and Fradd (1998) proposed "the notion of *instructional congruence* to indicate the process of mediating the nature of academic content with students' language and cultural experiences to make such content (e.g., science) accessible, meaningful, and relevant for diverse students" (p. 12). Moje, Collazo, Carrillo, and Marx, (2001) built on this concept of instructional congruence. The way students use scientific discourse is shaped by the everyday discourses that they bring to the classroom. To help students develop scientific discourse, teachers need to develop students' awareness of different discourses and make connections between students' everyday discourse and science discourse (Moje et al., 2001).

Focusing on science as a discourse with distinct language forms and ways of knowing, such as building theories, analyzing data, and communicating their findings, can help language-minority students learn to think and talk scientifically (Rosebery, Warren, & Conant, 1992). Students need to understand how constructing an explanation in science or supporting a claim in science looks different than in everyday life. Teachers also need to draw from students' everyday discourse (Moje et al., 2001) and make connections about the similarities between scientific discourse and everyday discourse. For example, a teacher may want to discuss how "using evidence" or "constructing an explanation" is similar and different in science compared with students' everyday lives.

Method

Instructional Context

This study occurred during a middle school chemistry unit, *How can I make new stuff from old stuff?* (*Stuff*) (McNeill et al., 2004), which we developed using a learning-goals-driven design model (Reiser, Krajcik, Moje, & Marx, 2003). The unit is contextualized in two everyday substances, soap and lard, with the students ultimately investigating how to make soap from lard. During the instructional sequence, students experience other phenomena as well, but they cycle back to soap and lard as they delve deeper into the different content learning goals. The learning-goals-driven design model emphasizes the alignment of the materials with national standards (AAAS, 1993; NRC, 1996). During the 8-week chemistry unit, students learn about substances and properties, chemical reactions, and conservation of mass, both at the phenomena level and the particulate level. Besides content learning goals, the unit also focuses on scientific inquiry practices. During the unit, students design investigations, conduct investigations, analyze data, create models, and construct scientific explanations. Frequently, the construction of scientific explanations is the culminating event in a lesson and supports the meaning-making by helping students apply scientific principles to understand the results of their investigations. For example, students conduct an investigation in which they combine a copper penny with acetic acid (i.e., vinegar). After they collect their data, they write a scientific explanation about whether or not a chemical reaction occurred.

When designing the *Stuff* curriculum materials, we incorporated educative components in the material including instructional practices to support students in writing explanations. By educative curriculum materials, we mean teacher materials that are specifically designed to promote teacher learning and practice (Ball & Cohen, 1996; Davis & Krajcik, 2005). We also provided teachers with professional development in which we discussed many aspects of the unit including how to support students in the construction of scientific explanations. Although we suggested that teachers use different instructional practices, we realized that the use and enactment of the curriculum materials would vary by teacher (Remillard, 2005). We view teaching as a design activity where teachers use their own resources and capacities to make meaning and adapt curriculum materials for their particular classrooms (Brown, 2004).

During the unit, the materials suggest that teachers introduce students to the concept of scientific explanations through a focal lesson. We were specifically interested in what instructional practices teachers used during this lesson, because of the lesson's explicit focus on explanation. This lesson occurs about 2 weeks into the unit after students have collected data for the various properties of lard and soap (i.e., color, hardness, solubility, melting point, and density). Initially, students gather data and write a scientific explanation using their prior understanding of explanation as a guide. After students write their explanations, the materials suggest that the teachers use a number of instructional practices to support students in scientific explanation. For example, they suggest that the teacher introduce the scientific explanation framework and define the three components (i.e., claim, evidence, and reasoning). The materials also suggest connecting the explanation framework to an everyday example, such as making the claim that an individual is the ''best singer'' or the ''best quarterback,'' and discussing how a scientific explanation is similar and different from an everyday example. Finally, the materials suggest that the teacher model how to construct a scientific explanation. The materials provide three hypothetical examples of weak and strong explanations that the teacher may use to model the use of the explanation framework to evaluate the explanations for the quality of the three components. After the discussion of scientific explanations, students then critique and rewrite their own explanations.

In the focal lesson on explanation, the instructional materials explicitly discuss three of the four instructional practices we are interested in investigating: defining, modeling, and connecting to everyday explanation. Unfortunately, in looking back at the lesson we found that although we discussed the rationale with the teachers, we did not suggest that they discuss it with their students. The curriculum materials discuss the rationale behind scientific explanation in an introductory section for teachers to help them understand what scientific explanation is and why it is important. However, the curriculum materials do not suggest that teachers discuss this rationale with their students. Nevertheless, we decided to still look for this strategy in the videotapes of the focal lesson to see if teachers were using it and whether it influenced student learning. Any findings about this and the other practices will inform future revision of the curriculum materials.

For students to learn how to evaluate data, they need numerous opportunities to evaluate rich, complex models of data (Chinn & Brewer, 2001; Lehrer & Schauble, 2002). We believe students also need numerous opportunities to engage in scientific explanations. After the focal lesson, students construct approximately ten scientific explanations during the unit.[1] Students record the results of their investigations and scientific explanations on student investigation sheets, which provide written scaffolds to support their explanation construction. These written scaffolds provide both context-specific and generic support and fade, or provide less detail, over time. The design was informed by research we conducted during the previous enactment in which we found that fading written support resulted in students constructing stronger explanations on the posttest, particularly for reasoning (McNeill et al., 2006).

Participants

Participants included 13 teachers out of a pool of 23 teachers involved in a larger curriculum materials development project. Many of the teachers had previously worked with the university researchers as part of research to develop project-based science curriculum. Teachers from the larger pool were asked to participate in this study due to the ability of researchers to get to the schools to collect data, and we wanted to include teachers from the different areas (e.g., large urban, small city, and suburban). Two of the teachers were men and the other 11 were women. Seven of the teachers were African American, five were Caucasian, and one chose not to report her ethnicity. All 13 teachers were certified to teach science although they had a wide range of experiences with teaching. The number of years they had taught science ranged from one teacher who was in her second year of teaching to another teacher who was in her thirtieth year of teaching. For the group of 13 teachers, the average number of years of experience teaching science was 10. Although there was a range of experiences in the group, all teachers were interested in using the inquiry-oriented curriculum materials; consequently, they may be more likely to support inquiry in their instruction than a random sample of middle school teachers.

The 13 teachers taught 1197 seventh grade students. Nine of the teachers (Teachers A–I) worked in a large midwestern urban area (Urban A) where the majority of the students were African American (over 90%) and from lower- to lower-middle-income families. Eight of these teachers worked in public schools and one taught in a charter school (Teacher I). Two teachers (Teachers J and K) were from an independent middle school in a small midwestern city (City B), where the majority of students were Caucasian and middle- to upper-middle-income families. One teacher (Teacher L) taught in a second, large midwestern urban area (Urban C) in a public school. The student population in this school was ethnically diverse (approximately 44% Caucasian, 34% African American, 12% Hispanic, and 10% Asian), with the majority of students from lower- and middle-income families. The last teacher (Teacher M) taught in a suburb of the second, large urban area (Suburb D). The student population in this school was ethnically diverse (approximately 45% Caucasian, 36% African American, 16% Hispanic, and 2% Asian) and the majority of these students were from lower- and middle-income families. As supported by the range of pretest scores at the beginning of the unit, students began the unit with a diversity of prior experiences and content knowledge. Unfortunately, our agreement with the schools only allows us to collect students' gender but not other demographic data from individual students.

Evaluating Teacher Instructional Practices

For each of the 13 teachers, we videotaped their enactment of the focal lesson on scientific explanation. We chose this lesson because we were specifically interested in how the teachers introduced scientific explanations to their students. Furthermore, as the lesson specifically focused on helping students understand how to construct scientific explanations, we believed there was a greater chance of observing the teachers' use of these different instructional practices during this lesson than any of the other lessons in the unit. To validate this assumption, we analyzed the videotapes for the next two lessons in which students wrote scientific explanations for a subset of the 13 teachers (four teachers), based on available videotape data. We found that the teachers used the four instructional practices, either in a similar manner or less frequently in these next two lessons, suggesting that the focal lesson was the appropriate choice for examining teacher instructional practices[2].

The focal lesson typically took teachers between 1 and 2 days, with one teacher requiring the beginning of a third day to complete the lesson. We developed the coding schemes from our theoretical framework, our experiences from a preliminary study (Lizotte et al., 2004), and an iterative analysis of the data (Miles & Huberman, 1994). Two of the codes, defining scientific explanation and modeling scientific explanation, were adapted from a preliminary study of videotape data from the previous enactment of the *Stuff* curriculum unit, which included six teachers (Lizotte et al., 2004). After finalizing the coding schemes, each lesson was scored by one of two raters. We randomly selected 46% of the lessons (6 of the 13 teachers) to be coded by the second independent rater. The interrater reliability, determined by percent agreement, was 82%. All disagreements were resolved through discussion. In what follows is a detailed description of the coding schemes for each of the four instructional practices.

To characterize how teachers defined scientific explanation, we gave each teacher a score from 0 to 5 for each of the three components of scientific explanation (see Table 1).

As accurate and complete definitions, we used the descriptions that were provided for the teachers in the curriculum materials. A claim is described as a statement or conclusion that answers the original question. Evidence is scientific data that is 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 and includes appropriate scientific principles. For each component, the curriculum offers different ways of discussing these definitions with students. The teachers did not need to use language identical to that in the curriculum, but the language did need to align with the intent of the materials. For example, one teacher received a Level 4 for the defining evidence code, because she talked about evidence as appropriate data, but did not discuss the idea of sufficiency. When the teacher asked the class to define evidence, students responded that evidence was "things that back up your claim" and "facts, numbers, or data." The teacher summarized the responses on the board by defining evidence as "data that supports or backs up the claim." This idea that data should support or back up the claim aligned with the idea of using appropriate data, but the teacher did not discuss the idea of sufficiency or including enough data.

In terms of providing a rationale, we were interested in whether teachers described a purpose for engaging in this scientific practice. Table 2 provides a description of the code.

Table 1
Code for defining scientific explanation

Level		Description of Level
0	Does not identify	The teacher did not mention the component during the focal lesson.
1	Incorrect definition	The teacher mentioned the component, but the definition of it was inaccurate.
2	No definition	The teacher mentioned the component, but did not explicitly define the component.
3	Vague definition	The teacher provided a vague definition of the component.
4	Accurate but incomplete definition	The teacher defined the component correctly, but the definition was incomplete. The definitions of claim, evidence, and reasoning each included two parts. Teachers who received this code only discussed one of the two parts.
5	Accurate and complete definition	The teacher provided a complete and accurate definition of the component, which included both parts.

Table 2
Code for making the rationale of scientific explanation explicit

	Level	Description of Level
0	Does not mention rationale	The teacher did not mention a rationale or a reason for creating an explanation during the focal lesson.
1	Vague rationale	The teacher mentioned a vague rationale, such as explanations being an important part of science or that scientists create explanations all the time.
2	Explicit rationale	The teacher explicitly mentioned the idea of constructing an explanation for an audience. Audience is discussed in terms of the purpose of an explanation is to convince or persuade someone else about the strength of a claim.

Specifically, we were interested in whether teachers discussed the importance of explaining phenomena and the idea of audience. Science is fundamentally about explaining phenomena and providing support for those explanations to justify them for other people, a particular audience. Merely providing a claim is not as convincing or persuasive as supporting that claim with evidence and reasoning. We gave each teacher a score of 0, 1, or 2 depending on their level of discussion around why people create explanations. For example, one teacher received a 1 for rationale, because she discussed the importance of explaining, but did not discuss the idea of audience. She told her class, “Explaining is probably the most important part of figuring out what is going on in science—it is what scientists do the most.”

For modeling scientific explanations, we examined how teachers used the three hypothetical examples in the curriculum about whether lard and soap are the same or different substances. To help students understand the quality of an explanation, a teacher needed to explicitly discuss the strengths and weaknesses of each component, rather than just provide an example or make a general statement that the explanation was good. We assigned each teacher a total of nine codes: claim, evidence, and reasoning codes for each of the three examples. Each code ranged from a Level 0 to Level 5. Table 3 provides a description of each level.

For example, one teacher received a Level 5 score for evidence when she modeled a weak explanation and then critiqued the evidence. The example stated, “The data table is my evidence that they are different substances.” After showing the example, the teacher asked the class “Did they give us some good evidence?” The class responded “No.” The teacher then critiqued this example, “What about the data table? I don’t know (gestures hands in the air). What on the data table? I don’t know (gestures hands in the air) . . . you did not give me any data to prove anything.” The teacher received a Level 5 for evidence because she explicitly identified the evidence as weak. Instead of providing a general statement about the data table, the explanation should have included specific data, such as the densities and melting points of fat and soap, to support the claim. After coding each example, we then averaged across examples to assign each teacher a mean score for each explanation component.

For connecting scientific explanation to everyday explanation, we were interested in whether teachers brought up everyday examples, such as art or sports, to help students understand how to construct explanations in science. Specifically, we were interested in whether teachers discussed how the general structure of a scientific explanation applied to everyday explanations. We provided each teacher with a rating of 0–4, depending on how many of the components they discussed in relation to an everyday example. The coding is described in Table 4.

Table 3
Code for modeling scientific explanation

	Level	Description of Level
0	Incorrect identification	The teacher incorrectly identified the component in the explanation. For instance, a teacher might say that an example does not include a claim when in fact it did include a claim.
1	Does not identify	The teacher did not mention whether the example included the component.
2	Identifies too much	The teacher identified more than the component in an explanation. For instance, a teacher might say that the claim in an example was ‘‘Fat and soap are different substances. Fat and soap have different colors. The second sentence is in fact part of the evidence so the teacher has identified more than the claim in this example. This score could only apply if the example included a component.
3	Vague identification	The teacher made a vague statement that an explanation did or did not include the component, but did not explicitly address why the example did or did not include that component. For instance, a teacher might simply say that an example includes reasoning without discussing where the reasoning is in the example or why it counts as reasoning.
4	Identifies too little	The teacher explicitly identified only a portion of a component. For instance, an example explanation may include three pieces of evidence and a teacher only discusses two of these pieces of evidence. A teacher could only receive this code if a component included multiple parts (e.g., three pieces of evidence).
5	Accurate and complete identification	The teacher explicitly identified the component in the example or explicitly stated that the explanation did not include that component.

One teacher received a Level 3 for her discussion of an art example. The teacher stated, ‘‘I am the best artist in this room—that is my claim . . . My evidence is that I understand what good art looks like. I draw beautiful stick figures and everyone understands my artwork when I put it on the board.’’ This teacher received a Level 3, because she explicitly talked about both the claim and evidence in her everyday example.

Table 4
Codes for connecting scientific explanation to everyday explanation

	Level	Description of Level
0	Does not mention an everyday example	The teacher does not mention an everyday example.
1	Discusses an everyday example, but not components	The teacher talks about an everyday example, such as basketball, tennis shoes, or allowances, but does not explicitly discuss the ideas of claim, evidence or reasoning in relation to the example.
2	Discusses an everyday example, including one component	The teacher talks about an everyday example, and explicitly talks about one of the three components (i.e., claim, evidence, or reasoning) in relation to the example.
3	Discusses an everyday example, including two components	The teacher talks about an everyday example, and explicitly talks about two of the three components (i.e., claim, evidence, or reasoning) in relation to the example.
4	Discusses an everyday example, including three components	The teacher talks about an everyday example, and explicitly talks about all three of the components (i.e., claim, evidence, or reasoning) in relation to the example.

Although these codes are far from exhaustive in terms of capturing what instructional practices the teachers used to support explanation, we were interested if they captured an aspect of teacher practice in the focal lesson that would predict student learning of scientific explanation.

Assessing Students' Explanations

To measure student learning, we collected pre- and posttest data. Students completed identical pre- and posttests that consisted of 15 multiple-choice items and four open-ended items. Three of the four open-ended items asked students to write scientific explanations for the three content areas in the unit: substances and properties, chemical reactions, and conservation of mass. The analysis for this study focused on these three scientific explanation items. Appendix A provides examples of two of the scientific explanation test items. Successfully writing a scientific explanation requires both an understanding of the science content and an understanding of how to write a scientific explanation (McNeill et al., 2006). Consequently, assessing students' scientific explanation should combine an analysis of the conceptual understanding and an analysis of the structure of the scientific explanation (Sandoval & Millwood, 2007). To assess students' scientific explanations, we adapted a base scientific explanation rubric for each of the three explanation tasks (see McNeill et al., 2006; McNeill & Krajcik, 2007). The base explanation rubric includes the three components that we then adapt to create a specific rubric to address the particular content and task. Appendix B provides an example of a specific rubric for the scientific explanation test question #1. A more complete description of our coding process and examples of student work can be found in the study by McNeill and Krajcik (2007).

Each explanation was scored by one rater. We then randomly sampled 20% of the tests, which were scored by a second independent rater. Our estimates of interrater reliability were calculated by percent agreement. Our interrater agreement was 97% for claim, 95% for evidence, and 97% for reasoning for the three explanations. Only students who completed both the pretest and posttest were included in the analysis. Because of high absenteeism in the urban schools only 835 students completed both the pre- and posttests. Of these students, 51% were female and 49% were male.

Results

Our analyses address the following three questions: (1) Did students' explanations improve from pre- to posttest and, if so, did this improvement vary by teacher? (2) What instructional practices did teachers engage in during the focal lesson to support students' explanations? (3) Did the teachers' instructional practices measured during the focal lesson predict student learning of scientific explanations?

Students' Pre- and Posttest Explanation Scores

We examined whether students' explanation scores improved significantly from the pre- to posttest. We summed students' explanation scores across the three explanation test items (substances, chemical reactions, and conservation of mass). We then analyzed their composite explanation score, which is a sum of their claim, evidence, and reasoning scores, as well as each component separately. Table 5 provides the results from this analysis.

Students showed significant learning gains on their composite explanation scores as well as on each separate component. This suggests that students became more adept at constructing scientific explanations during the instructional unit. Similar to our previous research (Lizotte et al., 2004; McNeill et al., 2006), we see that students have the most difficulty with the reasoning

Table 5
Overall student learning of scientific explanation (n = 835)

Score Type	Maximum	Pretest *M* (*SD*)	Posttest *M* (*SD*)	*t* (834)[a]	Effect Size[b]
Composite score	11.25	1.37 (1.48)	4.27 (2.48)	35.16***	1.96
Component					
Claim	3.75	0.87 (1.01)	2.05 (1.18)	25.54***	1.17
Evidence	3.75	0.42 (0.57)	1.28 (0.99)	24.86***	1.51
Reasoning	3.75	0.08 (0.25)	0.94 (0.94)	26.98***	3.44

[a]One-tailed, paired *t*-test.
[b]Effect size is the difference between pretest *M* and posttest *M* divided by pretest *SD*.
***$p < 0.001$.

component, but that the reasoning scores also demonstrate the greatest improvement from pre- to posttest as indicated by the greater effect size for reasoning compared with claim and evidence.

We also examined whether there was a significant difference in student learning between teachers. Figure 1 displays the effect sizes of the 13 teachers for students' total explanation scores.[3]

Although each teacher's students had significant learning gains for their explanations (p-values <0.001), the effect sizes ranged from 1.11 to 5.84. Teacher A is an exceptional teacher in Urban A. Because her effect size was much larger than that of the other teachers, we ran all of the future analyses in this study both with and without her students in the dataset. Including her students did not alter the significance level or direction of any of the results, so we included her in all analyses presented. In other work, we are looking at her case more closely to determine the unique characteristics of her classroom practice. We tested whether there was a significant teacher effect by performing an analysis of covariance (ANCOVA) on students' posttest explanation scores with the pretest explanation scores as the covariate and the teacher as the fixed factor. There was a significant teacher effect with the student learning gains of some teachers being greater than those of other teachers, $F(12, 821) = 16.429$, $p < 0.001$. There was also a significant interaction

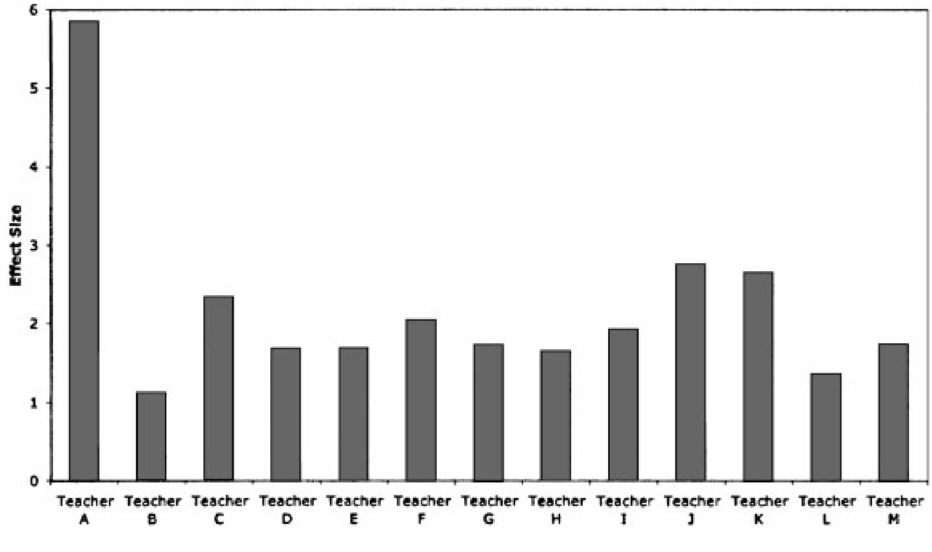

Figure 1. Effect size across teachers for students' scientific explanations.

between the teacher and students' pretest scores, $F(12, 821) = 2.776, p < 0.01$, suggesting that the effect of a student's pretest on his or her posttest varied by teacher. This suggests that there was differential learning of scientific explanation across teachers. Consequently, we were interested in whether teachers' practices during the focal lesson explained any of the between teacher variance.

Teachers' Practices During the Focal Lesson

We examined whether there was differential use of the four instructional practices during the focal lesson by the 13 teachers or if all of the teachers engaged in similar practices. Table 6 displays the descriptive statistics for each of the practices. For both defining and modeling, we created an overall composite score where we summed each teacher's scores for claim, evidence, and reasoning, and we also examined the scores for each component. We created a composite score to provide a more holistic view of the teachers' practices for explanation.

Defining Scientific Explanation

In terms of defining scientific explanation, all 13 teachers defined the different components, but the accuracy and completeness varied. Our coding scheme (see Table 1) included a scoring of just mentioning a component as a Level 2, a vague definition as a Level 3, and an accurate and complete definition as a Level 5. For all three components, the highest score was a Level 5, suggesting that at least one teacher accurately and completely defined the component. The lowest score for claim and evidence was a Level 2, which means that a teacher mentioned it without defining the component, and for reasoning it was a Level 0, which means one teacher never mentioned reasoning. Of the three components, claim had the highest average score, suggesting that the teachers most completely and accurately defined claim. There was the least variation in how the teachers discussed evidence, with most teachers referring to it vaguely as data, with little discussion of appropriateness or sufficiency. Finally, reasoning had the lowest average, but the greatest variation. Some teachers discussed extensively the idea of including a scientific principle to connect the claim and evidence, whereas one teacher did not even mention the concept of reasoning.

Making the Rationale of Scientific Explanation Explicit

When introducing scientific explanation, very few teachers discussed the rationale behind why an individual may want to construct a scientific explanation. Only 2 of the 13 teachers

Table 6
Descriptive statistics for teachers' instructional practices ($n = 13$ teachers)

	M (*SD*)
Defining scientific explanation	9.85 (2.38)
Defining claim	3.54 (1.05)
Defining evidence	3.31 (0.75)
Defining reasoning	3.00 (1.41)
Rationale of scientific explanation	0.15 (0.38)
Modeling scientific explanation	9.00 (2.92)
Modeling claim	3.21 (1.13)
Modeling evidence	3.26 (0.95)
Modeling reasoning	2.54 (1.24)
Connecting to everyday explanation	0.23 (0.44)

discussed a rationale behind completing explanation: Teacher J taught in City B and Teacher N taught in Urban C. Both teachers discussed the idea that the goal of science is to explain phenomena. Considering that the materials did not explicitly discuss this idea, it is not surprising that teachers did not include it in their classroom practice.

Modeling Scientific Explanation

All except one of the teachers modeled how to construct a scientific explanation. Of the other 12 teachers, there was a range of discussion in terms of the explicitness and accuracy of the three different components. Our coding scheme (see Table 3) included: not modeling a component as a Level 1; a vague identification of a component as a Level 3; and an accurate and complete identification of a component as a Level 5. For all three components, the teachers' scores ranged from Level 1 to Level 5. No teacher received a Level 0, which would have meant that a teacher incorrectly modeled a component. Similar to defining the three different components, the average claim score was again the highest, suggesting that the teachers accurately modeled how to construct a claim more than the other two components. Again, evidence had the least variation, with a mean score of 3.26, suggesting that the majority of teachers vaguely identified evidence when they modeled it for their students. Finally, similar to defining, reasoning again had the lowest mean and the most variation. The majority of teachers did not accurately and completely identify the reasoning, although there was one teacher who did receive a Level 5.

Connecting Scientific Explanations to Everyday Explanations

Similar to discussing the rationale behind scientific explanation, connecting scientific explanations to everyday explanations rarely occurred in the classrooms observed. Of the 13 teachers only 3 discussed everyday examples during the focal lesson. In two of the cases, the teachers discussed all three components in relation to an everyday example. In the third case, the teacher just discussed what a claim and evidence would look like in an everyday example (art).

Effect of Teachers' Instructional Practices on Students' Explanations

We created a hierarchical linear regression model to determine whether there was a relationship between teachers' instructional practices to support scientific explanation construction during the focal lesson and student learning of scientific explanations.[4] We were interested in whether the way they introduced scientific explanation to their students influenced students' learning of scientific explanation. We z-scored the outcome variable, the explanation posttest score, so that the unstandardized regression coefficients would be equivalent to the effect size. We also z-scored the pretest to keep it on the same scale. The rest of the variables we left in their original metric for ease of interpretation. We wanted to be able to talk about the effect of vaguely defining the different components of a scientific explanation, as compared with accurately and completely defining the components, not about a change in one standard deviation. Because teachers rarely completed both making the rationale explicit and connecting to the everyday, we could not treat these as continuous variables. Rather we dummy-coded both variables, so that the variable included in the regression model only indicated whether the teacher did (1) or did not (0) complete the practice. Because each of the 13 teachers received a distinct score for both defining and modeling at the composite level, we decided to treat these variables as continuous. Before conducting this analysis, we also created interaction terms.[5] We were interested in whether there was an interaction between the different teacher instructional practices. Did the effect of one teacher practice depend on the quality of a teacher's engagement in a second teacher practice? We calculated the interaction terms for the four teacher instructional practice predictors by multiplying each pair of predictors for a total of six interaction terms. The product term represents a combined effect of the two variables that is unique or goes above and beyond the separate additive effects of the two variables.

We used a hierarchical regression model because variables are grouped theoretically and then the groups are added one at a time (see Table 7). Model 1 includes the student measures from the beginning of the unit: gender and pretest score. Model 2 includes the four measures of teacher practices: defining scientific explanation, making the rationale of scientific explanation explicit, modeling scientific explanation, and connecting scientific explanation to everyday explanation. Finally, in Model 3 we added the interaction terms using stepwise regression. The interactions were added one at a time until they no longer significantly increased the proportion of the variance in the outcome variable explained by the model at an alpha level of 0.05. We used stepwise regression for the interaction terms because we only wanted to include the significant interactions in the regression model and we did not have a theoretical reason to include the interactions for some teacher practices and not others.

Table 7 includes the results for the regression analysis with the explanation posttest as the outcome variable, including the unstandardized regression coefficients and significant levels for each of the independent variables. The first group, which included gender and the pretest score, was significant for students' posttest, $F(2, 832) = 64.382$, $p < 0.001$. This regression model explained 13.4% of the variance in students' posttest scores. It is not surprising that students' performance on the pretest explained a large percentage of the variance on the posttest. In other words, students who scored higher on the pretest were more likely to score higher on the posttest. Gender was also significant, with females scoring higher on the posttest than males.

The change in the model resulting from the addition of the second group, which included the four instructional practices, was significant, $F(6, 828) = 43.560$, $p < 0.001$. Adding the teacher practices from the focal lesson explained 10.6% more of the variance in students' posttest explanation scores. Three of the four teacher practices significantly influenced student learning of scientific explanation: defining scientific explanation, making the rationale of scientific explanation explicit, and connecting scientific explanation to everyday explanation. Whether a

Table 7
Relationship between teachers' instructional practices and student learning (n = 835)

Independent Variables	Model 1: Student Measures	Model 2: Teacher Practices	Model 3: Interactions
Gender	0.163*	0.156*	0.148*
Pretest	0.355***	0.288***	0.266***
Defining scientific explanation		−0.124***	−0.147***
Rationale of scientific explanation		0.831***	0.548***
Modeling scientific explanation		0.011	0.007
Connecting to everyday explanation		−0.469***	−0.454***
Rationale × defining			0.407***
Constant	0.083	1.079***	1.338***
R^2	0.134***	0.240***	0.286***
Change in R^2		0.106***	0.045***

*$p < 0.05$; ***$p < 0.001$.

teacher discussed the rationale behind scientific explanation had the greatest effect. If a teacher did discuss the rationale compared to not discussing it, this resulted in significantly greater student learning of scientific explanation with an effect size of 0.831. Connecting scientific explanation to everyday explanation had a negative effect on student learning of scientific explanation. Furthermore, teachers who received a higher score for defining the components also had students with lower scientific explanation scores. Adding the four teacher practices also decreased the effect size of the pretest. This is important because it suggests that teacher practices can help students overcome some of their performance differences at the beginning of the unit.

The final model includes one significant interaction, rationale $\times$ defining.[6] The change in the model is significant for students' posttest scores, $F(7, 827) = 47.112$, $p < 0.001$, and explained 4.5% more of the variance. An interaction term suggests that the effect of one variable on student learning depends on another variable. Considering the importance of context in education, it is not surprising that the influence of one variable is going to depend on another variable. The final model explained a total of 28.6% of the variation in students' posttest explanation scores. This final model suggests the relative importance of each variable while considering the influence of the other variables.

In the final model, the direction and significance of the four main effects for the teacher practices remained the same. Explicitly discussing the rationale behind scientific explanation resulted in greater student learning gains, whereas defining the components of scientific explanation and linking scientific explanation to everyday explanation resulted in lower student gains. Modeling scientific explanations did not have a significant effect on student learning.

Besides the main effects, the interaction term suggests an interesting relationship between making the rationale explicit and defining explanation. Figure 2 depicts the interaction between explicitly discussing the rationale behind scientific explanation and defining the different components of explanation. The solid line in Figure 2 represents the effect size of posttest achievement when teachers provided students with the rationale behind scientific explanation, and the dotted line represents when the teachers did not provide the rationale. If a teacher discussed the rationale behind scientific explanation, then receiving a composite definition score of above 9 had a positive impact on student learning of scientific explanation. A composite score of 9 could be obtained by receiving a 3 for each of the components, which corresponds to providing vague definitions. A teacher who both discussed the rationale and accurately and completely defined the different components of scientific explanation had the greatest positive effect on students' posttest explanation achievement. If a teacher did not discuss the rationale behind scientific explanation, then accurately and completely defining the different components of an explanation actually had a negative impact on student learning. This suggests that the effect of one instructional practice may depend on the other practices that are a part of the classroom environment.

To summarize, the final model shows that three of the teacher practices used during the introduction to scientific explanation had a significant impact on student learning of scientific explanation. In terms of the main effects, discussing the rationale behind scientific explanation had a positive impact on student learning, whereas connecting scientific explanation to everyday explanation had a negative impact on student learning. The effect of defining the different components of explanation depended on whether or not a teacher also provided the rationale behind explanation. When a teacher provided the rationale, then defining the components had a positive impact, whereas, when a teacher did not provide a rationale, it had a negative impact. This suggests that the influence of an instructional practice depends on the context.

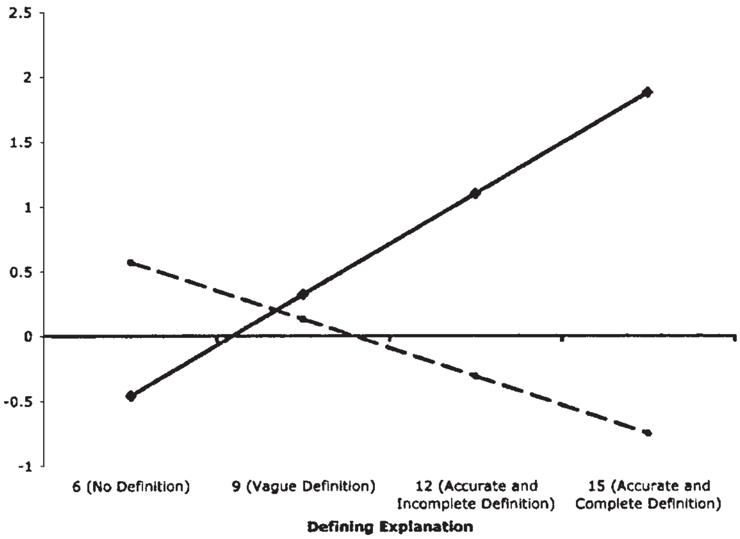

Figure 2. Interaction between rationale and defining (–◆–: Provided a Rationale; – –: Did not provide Rationale).

Discussion

The role of teachers is essential in supporting students in scientific inquiry practices (AAAS, 1993; NRC, 1996). Yet, like other researchers (Flick, 2000; Keys & Bryan, 2001), we argue that there have been few research studies that explicitly examine teacher instructional practices to support students in scientific inquiry. Specifically, we are interested in supporting students in constructing evidence-based scientific explanations, which are a fundamental aspect of scientific inquiry (Duschl & Osborne, 2002). Our results suggest that for all 13 teachers who completed the *Stuff* unit, their students had significant learning gains for scientific explanation. Yet teachers varied in their use of the instructional practices during the introduction of scientific explanation and this variation influenced student learning of scientific explanation.

Although all 13 teachers completed the same curricular unit, which explicitly focused on scientific explanation, we found that the adoption of the unit did not result in uniform instruction. The enactment of curriculum materials is a dynamic process mediated by a teacher's knowledge, beliefs and dispositions (Remillard, 2005). All teachers defined scientific explanations and the majority of teachers modeled how to construct scientific explanations, although the quality of these practices varied. For both defining and modeling, on average, teachers received the strongest scores for claim and weakest scores for reasoning, which also had the most variation. Because students received the least support from their teachers for reasoning, this is one explanation for why they had the most difficulty with this component. Few teachers discussed the rationale behind scientific explanations or connected scientific explanation to everyday explanation. Although there is a range of acceptable enactments of a curriculum, it is important for curriculum developers to clarify the essential components to help teachers in their adaptations (Remillard, 2005). By examining the effects of teachers' different adaptations, we hope to provide greater support to teachers.

We found that making the rationale of scientific explanation explicit for students during the introduction of scientific explanation resulted in greater student learning of scientific explanations. Instruction should help students understand the logic behind scientific inquiry practices (Kuhn et al., 2000). Helping students understand the rationale behind scientific explanations may help them see why they need to include evidence and reasoning to support their claims. When teachers include this instructional practice as part of their classroom instruction, students may obtain a stronger understanding of scientific explanation, which may help them in the construction of scientific explanations. Because so few teachers actually discussed the rationale behind scientific explanation in our study, we feel that to better understand this relationship we need to investigate more cases where the rationale is a part of classroom practice. In our regression model, we were only able to include the presence or absence of an accurate discussion of the rationale. It would be of interest to examine the depth of discussions around the rationale to see how this influences student learning. Based on the results of this study, we intend to revise the instructional materials to include an explicit discussion about the importance of discussing the rationale behind scientific explanation with students. Hopefully, by revising the curriculum materials, more teachers will engage in this instructional practice during their classroom instruction, allowing us to explore this practice in more depth.

Defining the different components of scientific explanations in the focal lesson increased student learning in some contexts, yet it decreased it in other contexts. There was an interaction between providing the rationale for scientific explanation and defining the different components of scientific explanation. When a teacher provided the rationale behind scientific explanation, then defining the different components resulted in greater student learning. However, when a teacher did not provide the rationale, then defining the different components of scientific explanation actually had a negative impact on student learning. Within classrooms, many factors influence student learning, including teachers, peers, and tools such as curriculum and software materials (Lampert, 2002). Tabak (2004) argued that students, tools, and teachers can act synergistically where they interact to support a specific learning goal. It is important to consider classrooms as complex systems when evaluating the effectiveness of any factor in terms of student learning. The results of our study suggest that even when looking at different teacher practices, it is important to consider what other practices occur within the classroom.

Previous research has found that being explicit about scientific inquiry practices (Herrenkohl et al., 1999) and providing students with different heuristics (Metz, 2000) can help students engage in scientific inquiry practices. Although providing students with a definition of scientific explanation and its components can help students engage in this practice, there is also the danger that explanation construction can become too algorithmic, formulaic, or procedural, without an understanding of the inquiry practice as a whole. We conjecture that, in classrooms where teachers focus on defining the parts, without a discussion of the rationale behind scientific explanation as a whole, constructing explanations becomes more algorithmic for students and they do not develop as deep an understanding of scientific explanation. Students may have understood scientific explanation as claim, evidence, and reasoning, but they did not understand the purpose behind the different components or how they fit together as a whole. Scientific explanations can become a rote task in which students do not understand why they are doing it or the motivation behind engaging in this complex practice (Kuhn & Reiser, 2005). This may explain why we found that when teachers defined the different components, but did not discuss the rationale, students had lower posttest explanation achievement. When supporting students in explanation and argumentation, it is important help motivate a ''need'' (Kuhn & Reiser, 2006) and help students understand the purpose behind this practice.

Modeling how to construct scientific explanations during the focal lesson did not significantly influence student learning of scientific explanations. Previous research has found that teacher modeling of scientific inquiry practices can encourage student success in these same practices (Crawford, 2000; Tabak, 2004). There are many possible reasons for why we did not find a significant effect in this study. For example, other teacher instructional practices may have had a stronger impact on student learning or it may have been more important to model scientific explanations over time and across different contexts than when explanations were first introduced.

The last instructional practice we examined was connecting scientific explanation to everyday explanation. We were surprised by the results of including this in our regression model. To help students develop a scientific discourse, teachers need to develop students' awareness of different discourses and make connections between students' everyday discourse and science discourse (Moje et al., 2001). Consequently, before conducting the analysis we thought that, if teachers made connections between everyday explanations and scientific explanations, then greater student learning would result. Our analysis suggests that the opposite occurred. Discussing everyday explanations in the classroom actually resulted in lower student posttest explanation achievement. Similar to our code for rationale, very few teachers engaged in this instructional practice and we were only able to include the strategy in our regression model in terms of the presence or absence. It may be that this negative effect is simply a result of our small sample of teachers who connected everyday explanation to scientific explanation. It is also possible that it is not the presence of this instructional practice that is important, but rather other characteristics of the instructional practice.

Our coding scheme (see Table 6) captured only whether or not teachers discussed an everyday explanation and what components they discussed. To further understand the effect of discussing everyday examples, we would need to examine more cases in which teachers used everyday explanations in their classrooms and assess the different ways they used the examples. In reexamining the three cases in which teachers discussed everyday examples, in all three instances the teachers discussed the similarities between everyday explanations and scientific explanations. To effectively use an everyday example, it may be more important to discuss the differences. Focusing on science as a discourse with distinct language forms and ways of knowing can help language-minority students learn to think and talk scientifically (Rosebery et al., 1992). Teachers need to discuss the differences between students' everyday discourses and scientific discourses (Lee, 2004). It may be that discussing everyday explanations is only helpful for students if it includes a discussion of the differences compared with scientific explanation, instead of discussing only the similarities, like the teachers in this study. To fully reveal the importance of this strategy, we would need to evaluate more teachers who compared everyday and scientific explanations in a variety of different ways.

The small sample size of our study, 13 teachers, may have influenced the results of our analysis, particularly for discussing the rationale behind scientific explanation and connecting scientific explanations to everyday explanations, because very few teachers used these instructional practices. Yet, we find the infrequency of these two practices and their possible influence on student learning to be important avenues for future research. In this study we also focused on the teachers' introduction of scientific explanation to their students during one lesson. Consequently, an extension of this study would be to track teacher instructional practices over time to see how their practices changed and how these practices influenced student learning. Future research should also look more closely at the interactions between the teacher and student, instead of simply focusing on the role of the teacher, as we did in this study.

Our study does provide some preliminary findings on how teacher practices can play an important role in student learning of scientific inquiry practices. Even when students are engaged in the same instructional unit, differential learning occurs that can be directly linked to instructional practices. Furthermore, the effect of these instructional practices can depend on the other supports available to students in the classroom. Developing a more in-depth understanding of these teacher practices is essential for supporting students in scientific inquiry practices, such as the construction of evidence-based scientific explanations.

Notes

[1]The number of explanations may vary slightly by teacher. There are optional lessons during the unit that teachers may choose to use with their students.

[2]We used a checklist for the four instructional practices to determine how frequently the teachers used these practices in the next two lessons. We analyzed one class period for Lesson 7 and two class periods for Lesson 8, for a total of 12 class periods across the four teachers. Two teachers engaged in similar practices in the next two lessons as compared with the focal lesson during which they both vaguely defined explanation and modeled explanations. The other two teachers provided slightly less support in the next two lessons. During the focal lesson, they both vaguely defined scientific explanation and modeled explanations. In the next two lessons, they again both vaguely defined scientific explanation, but did not model scientific explanation. Similar to the focal lesson, none of the four teachers discussed the rationale or connected them to everyday explanation in the next two lessons.

[3]We calculated effect size by dividing the difference between pretest and posttest means by the pretest standard deviation.

[4]Originally, we tried to run a hierarchical linear model (HLM), because were asking a multilevel question where students are nested in classrooms. Unfortunately, our sample size of teachers was not large enough to use HLM or to include the contextual factors in our statistical analysis. We did not have enough variance in students' learning of explanation between teachers.

[5]Before creating the interaction terms, we centered the two continuous variables, modeling and defining, to eliminate nonessential multicollinearity.

[6]For Model 3 and the previous two models we tested for multicollinearity by following the procedures of Cohen et al. (2003), wherein we examined the variance inflation factor (VIF), the tolerance, and the condition index. None of the independent variables in the final model or the two initial models exceeded their guidelines for VIF or tolerance. For the condition index, an index of >30 is considered a serious threat to multicollinearity, whereas an index of >15 indicates possible multicollinearity problems. For our final model, only two interactions were added stepwise to the model: rationale $\times$ defining and modeling $\times$ defining. Modeling $\times$ defining had a condition index of 16.529 and accounted for a sizable proportion of the variance for both the teacher practice of defining and the interaction between rationale and defining. Consequently, we were concerned about the collinearity of this independent variable with the other independent variables and chose to remove it from our final model.

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). Any opinions expressed in this work are those of the authors and do not necessarily represent either those of the funding agency, Boston College, or the University of Michigan. The authors thank all of the researchers involved with IQWST and CCMS, especially David Lizotte, Betsy Davis, Brian Reiser, and Leema Kuhn.

Appendix A: Examples of Scientific Explanation Test Questions

1. Carlos takes some measurements of two liquids—butanic acid and butanol. Then he stirs the two liquids together and heats them. After stirring and heating the liquids, they form two separate layers—Layer A and Layer B. Carlos uses an eyedropper to get a sample from each layer and takes some measurements of each sample. Here are his results:

Table A-1

	Measurements				
	Density	Melting Point	Mass	Volume	Solubility in Water
Before stirring and heating					
Butanic acid	0.96 g/cm^3	−7.9°C	9.78 g	10.18 cm^3	Yes
Butanol	0.81 g/cm^3	−89.5°C	8.22 g	10.15 cm^3	Yes
After stirring and heating					
Layer A	0.87 g/cm^3	−91.5°C	1.74 g	2.00 cm^3	No
Layer B	1.00 g/cm^3	0.0°C	2.00 g	2.00 cm^3	Yes

Write a **scientific explanation** that states whether a chemical reaction occurred when Carlos stirred and heated butanic acid and butanol.

2. Examine the following data table:

Table A-2

	Density	Color	Mass	Melting Point
Liquid 1	0.93 g/cm^3	No color	38 g	−98°C
Liquid 2	0.79 g/cm^3	No color	38 g	26°C
Liquid 3	13.6 g/cm^3	Silver	21 g	−39°C
Liquid 4	0.93 g/cm^3	No color	16 g	−98°C

Write a **scientific explanation** that states whether any of the liquids are the same substance.

Appendix B: Specific Explanation Rubric for Test Item #1 (Chemical Reactions)

Table B-1

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

References

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

Ball, D.L., & Cohen, D.K. (1996). Reform by the book: What is—or might be—the role of curriculum materials in teacher learning and instructional reform? Educational Researcher, 25, 6–8, 14.

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

Brown, M. (2004, April). Toward a theory of the design and use of curriculum materials. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

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

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

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

Cohen, J., Cohen, P., West, S.C., & Aiken, L.S. (2003). Applied multiple regression/correlation analysis for the behavioral sciences. Mahwah, NJ: Lawrence Erlbaum.

Crawford, B.A. (2000). Embracing the essence of inquiry: New roles for science teachers. Journal of Research in Science Teaching, 37, 916–937.

Crawford, T., Kelly, G., & Brown, C. (2000). Ways of knowing beyond facts and laws of science: An ethnographic investigation of student engagement in scientific practices. Journal of Research in Science Teaching, 37, 237–258.

Davis, E., & Krajcik, J. (2005). Designing educative curriculum materials to promote teacher learning. Educational Researcher, 34, 3–14.

Driver, R., Asoko, H., Leach, J. Mortimer, E., & Scott, P. (1994). Constructing scientific knowledge in the classroom. Educational Researcher, 23, 5–12.

Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. Science Education, 84, 287–312.

Duschl, R.A., & Osborne, J. (2002). Supporting and promoting argumentation discourse in science education. Studies in Science Education, 38, 39–72.

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, 915–933.

Flick, L.B. (1995) Complex instruction in complex classrooms: A synthesis of research on inquiry teaching methods and explicit teaching strategies. Paper presented at the annual meeting of the National Association for the Research in Science Teaching, San Francisco, CA.

Flick, L.B. (2000). Cognitive scaffolding that fosters scientific inquiry in middle level science. Journal of Science Teacher Education, 11, 109–129.

Fradd, S., & Lee, O. (1999). Teachers' roles in promoting science inquiry with students from diverse language backgrounds. Educational Researcher, 18, 14–20.

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, 451–493.

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.

Kelly, G.J., & Takao, A. (2002). Epistemic levels in argument: An analysis of university oceanography students' use of evidence in writing. Science Education, 86, 314–342.

Keys, C.W., & Bryan, L.A. (2001). Co-constructing inquiry-based science with teachers: Essential research for lasting reform. Journal of Research in Science Teaching, 38, 631–645.

Krajcik, J., Mamlok, R., & Hug, B. (2001). Modern content and the enterprise of science: Science education in the twentieth century. In L. Corno (Ed.), Education across a century: The centennial volume (pp. 205–231). Chicago, IL: NSSE.

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

Kuhn, L., & Reiser, B. (2005, April). 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.

Kuhn, L., & Reiser, B. (2006, April). Structuring activities to foster argumentative discourse. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

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

Lampert, M. (2002). Appreciating the complexity of teaching and learning in school: A commentary on Cobb; Froman and Ansell; McClain; Saxe; Schliemann; and Sfard. The Journal of the Learning Sciences, 11, 365–368.

Lee, O. (2004). Teacher change in beliefs and practices in science and literacy instruction with English language learners. Journal of Research in Science Teaching, 41, 65–93.

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

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.

Lehrer, R., & Schauble, L. (Eds.). (2002). Investigating real data in the classroom: Expanding children's understanding of math and science. New York: Teachers College Press.

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.

Marx, R.W., Blumenfeld, P.C. Krajcik, J.S., & Soloway, E. (1997). Enacting project-based science. The Elementary School Journal, 97, 341–358.

McNeill, K.L., & Krajcik, J. (2007). Middle school students' use of appropriate and inappropriate evidence in writing scientific explanations. In M. Lovett & P. Shah (Eds.), Thinking with data: Proceedings of the 33rd Carnegie Symposium on Cognition (pp. 233–265). New York: Taylor & Francis.

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., Lizotte, D.J., Krajcik, J., & Marx, R.W. (2006). Supporting students' construction of scientific explanations by fading scaffolds in instructional materials. The Journal of the Learning Sciences, 15, 153–191.

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.

Meyer, K., & Woodruff, E. (1997). Consensually driven explanation in science teaching. Science Education, 80, 173–192.

Miles, M., & Huberman, A.M. (1994). Qualitative data analysis: An expanded sourcebook. (2nd ed.). Thousand Oaks, CA: Sage.

Moje, E., Collazo, T. Carrillo, R., & Marx, R. (2001). ‘‘Maestro, what is quality?’’ language, literacy, and discourse in project-based science. Journal of Research in Science Teaching, 38, 469–498.

Moje, E.B., Peek-Brown, D., Sutherland, L.M., Marx, R.W. Blumenfeld, P., & Krajcik, J. (2004). Explaining explanations: Developing scientific literacy in middle-school project-based science reforms. In D. Strickland & D.E. Alvermann (Eds.). Bridging the gap: Improving literacy learning for preadolescent and adolescent learners in grades. 4–12. New York: Teachers College Press.

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

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

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: Lawrence Erlbaum.

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.

Remillard, J.T. (2005). Examining key concepts in research on teachers' use of mathematics curricula. Review of Educational Research, 75, 211–246.

Rosebery, A., Warren, B., & Conant, F., (1992). Appropriating scientific discourse: Findings from language minority classrooms. The Journal of the Learning Sciences, 2, 61–94.

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

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

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

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

Tabak, I., & Reiser, B.J. (1997) Complementary roles of software-based scaffolding and teacher–student interactions in inquiry learning. In R. Hall, N. Miyake, & N. Enyedy (Eds.). Proceedings of Computer Support for Collaborative Learning '97 (pp. 289–298). Toronto, Canada.

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, Walton, Willard, Woods, & Zarefsky. (1996). Fundamentals of argumentation theory: A handbook of historical backgrounds and contemporary developments. Mahwah, NJ: Lawrence Erlbaum.

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, 437–465.

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