

University of Nebraska - Lincoln
DigitalCommons@University of Nebraska - Lincoln

Papers in Natural Resources

Natural Resources, School of

2013

Fostering Student Sense Making in Elementary Science Learning Environments: Elementary Teachers' Use of Science Curriculum Materials to Promote Explanation Construction

Laura Zangori

University of Missouri, zangoril@missouri.edu


Cory Forbes

University of Nebraska-Lincoln, cforbes3@unl.edu

Mandy Biggers

The Pennsylvania State University

Follow this and additional works at: <http://digitalcommons.unl.edu/natrespapers>

 Part of the [Curriculum and Instruction Commons](#), [Elementary Education Commons](#), [Natural Resources and Conservation Commons](#), [Natural Resources Management and Policy Commons](#), [Other Environmental Sciences Commons](#), and the [Science and Mathematics Education Commons](#)

Zangori, Laura; Forbes, Cory; and Biggers, Mandy, "Fostering Student Sense Making in Elementary Science Learning Environments: Elementary Teachers' Use of Science Curriculum Materials to Promote Explanation Construction" (2013). *Papers in Natural Resources*. 748.

<http://digitalcommons.unl.edu/natrespapers/748>

This Article is brought to you for free and open access by the Natural Resources, School of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Papers in Natural Resources by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Published in *Journal of Research in Science Teaching* 50:8 (2013), pp 989–1017.

doi 10.1002/tea.21104

Copyright © 2013 Wiley Periodicals, Inc. Used by permission.

Submitted 2 July 2012; accepted 21 July 2013; published 13 August 2013.

An earlier version of this article was presented at the 2013 meeting of the American Educational Research Association in San Francisco, CA.

Fostering Student Sense Making in Elementary Science Learning Environments: Elementary Teachers' Use of Science Curriculum Materials to Promote Explanation Construction

Laura Zangori,¹ Cory T. Forbes,² and Mandy Biggers³

¹ Science Education, Department of Teaching and Learning, University of Iowa,
S113 Lindquist Center, Iowa City, Iowa

² School of Natural Resources, College of Agriculture and Natural Resources,
University of Nebraska–Lincoln, Lincoln, Nebraska

³ Department of Curriculum and Instruction, College of Education,
Pennsylvania State University, University Park, Pennsylvania

Corresponding author – L. Zangori; email laurazangori@gmail.com

Laura Zangori and Cory T. Forbes share first authorship. Mandy Biggers is second author.

Abstract

While research has shown that elementary (K-5) students are capable of engaging in the scientific practice of explanation construction, commonly-used elementary science curriculum materials may not always afford them opportunities to do so. As a result, elementary teachers must often adapt their science curriculum materials to better support students' explanation construction and foster student sense making. However, little research has been conducted to explore *if* and, if so, *how* and *why*, elementary teachers modify science curriculum materials to engage students in explanation construction. We use an embedded mixed methods research design to explore elementary teachers' ($n = 45$) curricular adaptations and pedagogical reasoning. We collected and quantitatively analyzed

a matched set of 121 elementary science lesson plans and video recorded lesson enactments to investigate the extent to which in-service elementary teachers engage in instruction to more productively support students' explanation construction. Our findings suggest that the curriculum materials heavily emphasized hands-on engagement and data collection over explanation construction and that the teachers' adaptations did not fundamentally alter scientific sense-making opportunities afforded students in the lesson plans. Interviews and other artifacts were also collected and analyzed to construct a multiple-case study of four of these elementary teachers. Findings from the case study suggest that the teachers' conceptions of explanation construction and concerns about the abilities of their students to engage in scientific explanations impacted their curricular adaptations.

Keywords: elementary science, elementary teachers, explanations, curriculum materials

Elementary students, like middle-school and secondary students, should be afforded opportunities to engage in scientific practices to develop deep conceptual understanding of natural phenomena and experience the ways in which scientific knowledge is generated (Duschl, Schwein-gruber, & Schouse, 2007). A crucial scientific practice is that of explanation construction (National Research Council [NRC], 2000, 2012), which requires students to give priority to evidence in formulating evidence-based explanations that answer an investigation question and build upon their knowledge (NRC, 2000). While the science education and learning sciences communities recognize numerous perspectives on the nature of evidence-based explanations (e.g., Braaten & Windschitl, 2011; McNeill & Krajcik, 2008; Osborne & Patterson, 2011), there is consensus that the link between evidence and explanation is central to scientific sense making (Duschl, 2008; NRC, 2012). Sense making in science rests upon fundamental epistemological assumptions about the validity of evidence, its use to ground claims, and the cultural norms that shape the negotiation of meaning around competing explanations. Most importantly, opportunities to construct evidence-based explanations have been found to be a strong predictor of student learning (McNeill & Krajcik, 2008; Songer & Wenk Gotwals, 2012).

A burgeoning literature base has demonstrated that elementary students can successfully engage in explanation construction and other scientific practices that comprise scientific sense making (Cavagnetto, Hand, & Norton-Meier, 2010; Hapgood, Magnusson, & Palincsar, 2004; Hardy, Jonen, Möller, & Stern, 2006; McNeill, 2011; Metz, 2011; Samarapungavan,

Mantzicopoulos, & Patrick, 2008; Songer & Wenk Gotwals, 2012) when using specialty-designed elementary science curriculum materials associated with specific research and development projects. However, recent evidence shows that evidence-based explanations are frequently underemphasized in elementary science learning environments (Forbes, Biggers, & Zangori, 2013; Minogue, Madden, Bedward, Wiebe, & Carter, 2010) where widely available, commercially published elementary science curriculum materials are implemented. Consistent with theoretical perspectives on the teacher–curriculum relationship (Forbes & Davis, 2010; Remillard, 2005), this trend may be attributed not only to the elementary science curriculum materials that teachers use (Biggers, Forbes, & Zangori, in press), but also *how* these resources are used in light of teachers’ knowledge, beliefs, or priorities regarding explanation construction (Beyer & Davis, 2008; Forbes et al., 2013; Minogue et al., 2010).

However, few studies have explored how in-service elementary teachers use science curriculum materials to afford students opportunities to engage with scientific explanations. In this embedded mixed methods study (Creswell & Plano Clark, 2011), we address this gap in the literature by investigating opportunities elementary students are afforded to engage in explanation construction in both planned and enacted science instruction and elementary teachers’ instructional practices and underlying reasoning about promoting and supporting students’ explanation construction. We have chosen to focus on two features of inquiry (NRC, 2000)— *giving priority to evidence* and *formulating evidence-based explanations*—because our prior research has shown that elementary science underemphasizes scientific sense making (Biggers et al., in press; Forbes & Davis, 2010; Forbes et al., 2013; Zangori & Forbes, 2013). In this study, we examine the extent to which these features of inquiry are present in the original, unmodified curriculum materials (identified here as “lesson plans” or “planned lessons”) and teachers’ lesson enactments (identified here as “lesson enactments” or “enacted lessons”). Our research questions are:

1. How and to what extent does elementary teachers’ planned and enacted science instruction engage students in *giving priority to evidence* and *formulating evidence-based explanations*?
2. How does in-service elementary teachers’ pedagogical reasoning about *giving priority to evidence* and *formulating evidence-based explanations* inform the ways in which they engage students in explanation construction in the classroom?

Background and Theoretical Framework

Explanations as Cause, Effect, and Mechanism

While the importance of scientific explanations in elementary science learning environments is highlighted in science education reform efforts (Duschl et al., 2007; NRC, 2000, 2012), many questions remain as to what constitutes a scientific explanation, how students should construct them in the classroom, and how teachers should scaffold students to do so (Braaten & Windschitl, 2011; McNeill & Krajcik, 2008; Osborne & Patterson, 2011). Multiple views of explanation construction have been articulated in the field. Some have foregrounded the scientific reasoning that connects claims and evidence (Beyer & Davis, 2008; McNeill & Krajcik, 2008). Others have highlighted the centrality of scientific questions as driving the formulation of claims and mobilization of evidence (Cavagnetto et al., 2010). Still others have operationalized explanation construction in terms of students' ability to attribute unobservable, underlying mechanisms to observable phenomena (Braaten & Windschitl, 2011; Chinn & Malhotra, 2002; Windschitl, Thompson, & Braaten, 2008). These perspectives also differ in terms of how integrated explanation construction and other scientific practices, particularly argumentation, are perceived to be. While some delineate a difference between explanations and argumentation (e.g., Berland & Reiser, 2008; Braaten & Windschitl, 2011; Osborne & Patterson, 2011) other research has defined explanation construction within an argumentation framework (e.g., Berland & McNeill, 2012; Cavagnetto et al., 2010; McNeill & Krajcik, 2008; Ruiz-Primo, Li, Tsai, & Schneider, 2010).

Both explanation and argument are critical scientific practices (NRC, 2012). Consistent with Osborne's perspective, we note the purpose of an explanation as to "offer a plausible causal mechanism" while the goal of argumentation is to "persuade" others (Osborne & Patterson, 2011, p. 8). Our prior works finds that while elementary curriculum materials and teachers' enacted instruction most frequently provides opportunities for students to engage in questioning and data collection, the materials and enactments rarely provide opportunities to engage in the formulation and comparison of evidence-based explanations (Biggers et al., in press; Forbes & Davis, 2010; Forbes et al., 2013; Zangori & Forbes, 2013). Others (e.g., Berland & McNeill, 2012; McNeill & Krajcik, 2008) have identified scientific explanation construction as part of an argumentation framework to attempt to incorporate both practices into instruction. However, while instruction should incorporate both explanation and

argumentation, it follows logically that such instruction is not possible if the core component—the scientific explanation—is absent. In short, explanations may be formulated without any direct attempt to persuade but scientific argument cannot occur without explanations.

To examine the prevalence of scientific explanations in the classroom, we have closely aligned our definition of scientific explanations with Braaten & Windschitl's (2011) perspective grounded in cause and effect. We have built upon the Windschitl framework to clarify and operationalize “mechanisms”—what Braaten and Windschitl (2011) refer to as “unobservable, theoretical components” (p. 662). Mechanisms provide reasons by which a cause can bring about an effect (NRC, 2012). A scientific explanation is distinct from a scientific “explication” (Braaten & Windschitl, 2011, p. 651). Scientific explications are descriptions of “what” happened during the lesson (e.g., which plant grew taller or how long before butterflies emerged from the chrysalis). These explications are important in classroom science because they help students clarify their evidence (cause and effect). However explications alone do not identify mechanisms. An explanation occurs when students are able to build on their existing knowledge to understand “how” and “why” they observed what they did (e.g., the mechanism). Within this view, the purpose of explanation construction in the science classroom is for students to make sense of how the world works by connecting the cause and effect of natural phenomena (data and evidence) with its underlying mechanism (explanation).

Explanation construction in elementary science learning environments is also viewed as a “pragmatic” (Braaten & Windschitl, 2011, p. 644) scientific practice in that the constructed explanations are dependent on the norms of the classroom, the evidence that the students derive from their classroom inquiry, and the mechanism that is intentionally targeted through curriculum and instruction (Braaten & Windschitl, 2011; Magnusson & Palincsar, 2005; Salmon, 1998). If students are not afforded opportunities to establish connections between cause, effect, and mechanism, their explanations may be limited to what they have observed, but the phenomenon will remain as a “black box whose internal workings are mysterious” (Salmon, 1998, p. 89) and students will be unable to *make sense* of the phenomenon (Braaten & Windschitl, 2011; Chinn & Malhotra, 2002; Duschl et al., 2007; NRC, 2012; Salmon, 1998; Trout, 2002). We define student opportunities to make sense of the phenomenon as scientific sense making which is the student's opportunity to connect cause and effect with the underlying mechanism (Berland & Reiser, 2008).

Scientific Explanations and Classroom Inquiry

Braaten and Windschitl's (2011) perspective on explanation construction aligns with theoretical perspectives on classroom inquiry articulated in science education reform that ground our work (Biggers et al., in press; Forbes et al., 2013; Forbes & Davis, 2010; Zangori & Forbes, 2013; Zangori, Forbes, & Biggers, 2012). It highlights two components—evidence and explanation—as core elements that are fundamental to scientific sense making (Duschl, 2008; Duschl et al., 2007; Magnusson & Palincsar, 2005). The interpretation of data and mobilization of evidence are crucial to establish causal accounts and the mechanism for *why* something occurred. This crucial relationship is clearly represented in Duschl's (2008) evidence explanation (E-E) model and the five essential features of inquiry framework (NRC, 2000) that highlight the relationship between evidence and explanation. Both models emphasize the need for students to be afforded opportunities to engage in scientific sense making through (a) selecting or generating data to become evidence (cause), (b) using evidence to ascertain patterns of evidence and models (effect), and (c) employing the models and patterns to propose explanations (mechanism). The first two of these activities embody feature of inquiry *giving priority to evidence* while the third is synonymous with *formulating evidence-based explanations* (NRC, 2000).

We rely upon both the NRC's (2000) five essential features of inquiry, Duschl (2008) E-E model, and the Braaten and Windschitl's (2011) perspectives on explanations to define and operationalize the components of cause, effect, and mechanism into discernible constructs that we can measure in the elementary classroom. First, we define and operationalize cause and effect. Students must engage with and represent real-world phenomena in the form of data and evidence (NRC, 2000, 2012). As Braaten and Windschitl (2011) suggest, data and evidence occur when a student “describes, summarizes, or restates a pattern or trend in data without making a connection to any unobservable/theoretical components” (p. 662). This aligns with Duschl's (2008) step a (selecting or generating data to become evidence) and step b (using evidence to ascertain patterns of evidence and models). At this stage, students are “explicating” their observations. We build upon these definitions offered by Braaten and Windschitl (2011) and Duschl (2008) through the essential features of inquiry (NRC, 2000) and scientific practices (NRC, 2012) to identify core tasks (Forbes et al., 2013) through which students establish and describe cause and effect for natural phenomena. First, students engage with the phenomenon of interest, whether it is through hands-on

practical activities, engaging with texts about the phenomenon, observing demonstrations, etc. Second, they work with data by, for example, recording observations and/or measurements. Third, they should organize and analyze data through graphing, quantitative transformations, and/or categorization so as to ascertain patterns or trends in observed phenomena. Finally, fourth, they should reflect upon and verify their data collection processes (i.e., reading a balance correctly or determining if an object sank or float) and data analysis strategies (accuracy of graphed data or double-checking categorization criteria). Engagement in these classroom tasks affords students opportunities to *explicate* cause and effect, or descriptions of “what” happened (Braaten & Windschitl, 2011).

Second, we define and operationalize mechanism from Braaten and Windschitl (2011) as well as Duschl (2008) step c (employing the models and patterns to propose explanations). At this stage, students are proposing the how and why—the mechanism—for their observations. To formulate scientific explanations for observed cause and effect, students must articulate a mechanism that describes “how” and/or “why” the phenomenon occurs. The “how” or “why” is what differentiates a description of an observed phenomenon and an explanation for it (Osborne & Patterson, 2011; Salmon, 1998; Trout, 2002). As Braaten and Windschitl (2011) suggest, explanation construction involves students articulating “a full causal story for why a phenomenon occurred” and using “unobservable/theoretical components of a model to explain an observable event/phenomenon” (p. 662). The “unobservable/theoretical components” of the model are the mechanism and the process accounting for these elements is, as Salmon (1998) described, the opening of black boxes. Again, building upon definitions of essential features of inquiry (NRC, 2000) and scientific practices (NRC, 2012) in the field, we identify core features of scientific explanations (Forbes et al., 2013) through which students describe how and why phenomena occur. Scientific explanations must (a) be supported by evidence; (b) answer a question driving the investigation; (c) be based upon students’ pre-existing ideas; and, (d) propose new understanding about the observed phenomenon.

Explanation Construction in Elementary Science Learning Environments

Despite the emphasis on explanation construction in science education reform (Duschl et al., 2007; NRC, 2012) and a growing body of evidence that elementary students can engage productively in a variety of scientific practices (Cavagnetto et al., 2010; Hapgood et al., 2004; Hardy et al., 2006; McNeill, 2011; Metz, 2011; Samarapungavan et al., 2008; Songer &

Wenk Gotwals, 2012), elementary students are often not afforded opportunities to engage in substantive sense making about science, including the formulation of evidence-based explanations (Beyer & Davis, 2008; Biggers et al., in press; Forbes et al., 2013; Metz, 2009; Minogue et al., 2010). Instead, elementary science is often characterized by an emphasis on hands-on activities designed to be engaging and motivating for students but that do not foreground scientific sense making (Metz, 1995). While no comprehensive review of elementary science curriculum materials has yet been conducted (Kesidou & Rosemann, 2002), some evidence suggests commonly-available, widely-used elementary science curriculum materials also more heavily emphasize students' engagement with phenomena than scientific sense-making practices (Biggers et al., in press; Forbes & Davis, 2010; Forbes et al., 2013). As a result, elementary teachers often must modify the curriculum materials they use to engage students in scientific practices such as explanation construction, which must be supported over time through instruction (Avraamidou & Zembal-Saul, 2005; Beyer & Davis, 2008). This assumption is grounded in broader perspectives on the teacher-curriculum relationship in which teachers' flexible use of curricular resources is an expected professional practice (Remillard, 2005).

However, the teacher-curriculum relationship is heavily influenced by the reasoning tools teachers leverage to make decisions about instruction. A body of research has illustrated elementary teachers' interactions with scientific practices and features of inquiry, including explanation construction. This research suggests that elementary teachers may not have a strong understanding on what constitutes scientific reasoning (Avraamidou & Zembal-Saul, 2005; Beyer & Davis, 2008), skip components of scientific explanation construction when it is included in the curriculum materials (McNeill & Krajcik, 2008; Ruiz-Primo et al., 2010), or assume elementary students are not able to engage in scientific reasoning (Metz, 2009). It is critical that elementary teachers develop robust knowledge of explanation construction and student reasoning so as to be able to use their curricular resources and scaffold students' efforts to formulate evidence-based explanations. Such scaffolding has been shown to be crucial to support early learners' sense making (Hapgood et al., 2004; Hardy et al., 2006; McNeill, 2011; Metz, 2011; Samarapungavan et al., 2008; Songer & Wenk Gotwals, 2012). Furthermore, when teachers explicitly highlight and scaffold students' explanation construction, it can lead to greater learning gains (McNeill & Krajcik, 2008; Ruiz-Primo et al., 2010; Songer & Wenk Gotwals, 2012).

Method

In this embedded mixed methods study (Creswell & Plano Clark, 2011); we analyze evidence of planned and enacted elementary science instruction to investigate the extent to which science curriculum materials and classroom instruction engage students in the scientific practice of explanation construction. We also conducted a multiple-case study (Yin, 2009) of four elementary teachers to investigate *how* and *why* they use their elementary science curriculum materials to engage and support students' explanation construction. The purpose of the multiple-case study is to elaborate and enhance the findings from quantitative analyses of teachers' planned and enacted science instruction.

Study Context and Participants

This empirical study is part of the Promoting Inquiry-Based Elementary Science through Collaborative Curriculum Co-Construction (PIESC3) project, a 3-year research and development effort grounded in a science professional development program for elementary teachers in a large, urban school district (Biggers et al., in press; Forbes et al., 2013; Zangori et al., 2012). The program's purpose is to support elementary teachers in evaluating, planning, and adapting their science curriculum materials to better engage students in scientific practices and inquiry (NRC, 2000, 2012). Project participants included 45 in-service elementary teachers from the partner district and four surrounding districts within a Midwestern State. The partner district is the second largest school district in this state and includes schools in both urban to rural communities. This district has a total of 16,000 students with 17 elementary schools. Out of these 17 elementary schools, teachers from 11 of these elementary schools participated in this study. These 11 elementary schools had a free/reduced lunch range of 18–83% while the district as a whole has a 63% free/reduced lunch. Teachers from the four surrounding districts included both rural and urban schools with a free/reduced lunch rate ranging from 3% to 89%. All participating teachers use common, widely-available, research-based, reform-oriented, kit-based elementary science curriculum materials from major commercial publishers (e.g., FOSS, STC, Insights). Of the 45 teachers involved in the project, eight from the partner district volunteered to be case study participants. Participation in this project was voluntary and all teachers were compensated for their involvement. The data presented here is from the first year of the project prior to the implementation of the program and serves as baseline data for the teachers' normal implementation of their curriculum materials.

Teachers' Planned and Enacted Science Instruction

Data Collection

All 45 teachers in the project were asked to provide three matched sets of original lesson plans and video recorded lesson enactments from a science unit of their choosing. Non-case study teachers used project-provided video recording equipment to self-record their three science lessons (all participating teachers were provided instructions for how to use the video-recording equipment). Secure digital (SD) memory cards were used for all digital video recordings and were submitted to the project team in the mail along with hard copies of the accompanying lesson plans from unit-specific teachers' manuals. Live observations were conducted and video recorded by the authors for each of the case study teachers' three lessons.

By the end of the academic year, 89.6% of participating teachers submitted matched sets of (a) original unmodified lesson plans (from published teacher manuals) and (b) video recorded lesson enactments (their teaching of these lessons). This resulted in 40 teachers each submitting lesson plans and video-recordings of lesson enactments for three lessons and one study teacher only submitted one matched lesson plan with video recorded enactment. Therefore, our total sample size was 121 matched sets of lesson plans and video recorded lessons ([40 teachers \times 3 lessons] + [1 teacher \times 1 lesson]). Throughout the article we refer to science lessons represented in the teachers' original curriculum materials as "lesson plans" and "planned lessons" and to their teaching of these lessons as "lesson enactments" and "enacted lessons."

Data Analysis

All 121 matched lesson plans and video recorded enacted lessons were scored using a newly-developed Practices of Science Observation Protocol (P-SOP—Forbes et al., 2013). The 20-item P-SOP is designed to provide a measure of the five essential features of inquiry (NRC, 2000) in classroom settings, which include learners (a) engaging in scientifically oriented questions; (b) giving priority to evidence; (c) formulating explanations from evidence to address scientifically oriented questions; (d) evaluating their explanations in light of alternative explanations, particularly those reflecting scientific understanding; and (e) communicating and justifying their proposed explanations. The P-SOP is designed to measure the classroom environment overall and is not a specific measure of either students or teachers but rather measures the presence of opportunities for engagement in the features of inquiry collaboratively.

For this study, our observational focus was on the presence of opportunities provided during the lesson for students to engage in the features of evidence and construct evidence-based explanations. Each item is scored on a scale of 0–3. Sets of four instrument items are grouped to provide submeasures of each of the five essential features of inquiry. A score of 0 indicates the absence of the task or practice an item is intended to measure. A score of 3 indicates observed classroom activity provided evidence of all aspects of the task or practice an item was intended to measure. Aggregate score for observed science instruction could range from 0 to 60, while the range of subscores for each feature of inquiry is 0–12. This scoring yielded two sets of scores: one for the lesson plans and another for the video-recorded enacted lessons.

Findings from field-testing of the P-SOP, in which it was shown to be valid and reliable for use in elementary settings, have been published elsewhere (Forbes et al., 2013) but are summarized again here. The instrument is designed to provide a measure of classroom inquiry as represented in the five essential features framework articulated by the NRC (2000). Instrument development was grounded in a full literature review and draft versions of the instrument were submitted for multiple rounds of external review by experts in the field. The pilot version was used by a team of two scorers to jointly score 124 elementary science lessons collected to evaluate the P-SOP. This entire sample was used to establish interrater coding reliability (Forbes et al., 2013). Intraclass correlation coefficients for inter-scorer reliability ranged from 0.76 to 0.9 across the entire instrument and each of the five inquiry feature submeasures. To establish instrument reliability, Cronbach's α values were calculated for the aggregate data; as well the five individual feature subscores, and ranged from 0.71 to 0.98.

In this study, we focused our analysis on two of these five features of inquiry measured in the instrument: *giving priority to evidence* and *formulating explanations from evidence*. Descriptions of constituent items for these two features of inquiry, as well as psychometric properties of each, are shown in **Table 1**.

In the instrument-development study (Forbes et al., 2013), internal, instrument reliability ranged from 0.68 to 0.93 for the eight items presented in Table 1. Item scores for these two features of inquiry were imported into SPSS for statistical analysis, the purpose of which was to quantitatively analyze feature submeasure scores and make comparisons between the original lesson plans and video recorded observations. Because multiple lesson plans and enactments occur for each teacher in a sequential manner, our analysis required that we account for both

Table 1. Descriptions and psychometric properties for features and items for students' giving priority to evidence and formulating evidence-based explanations

<i>Inquiry Feature</i>	<i>Feature α</i>	<i>Feature ICC</i>	<i>Item and Description</i>	<i>Item α</i>	<i>Item ICC</i>
Giving priority to evidence	0.71	0.53 ^a	Students engage with the phenomena targeted in the lesson's learning goals and performances through investigations, modeling activities, etc.	0.84	0.78 ^b
			Students work with data that are related to the phenomena of interest. Data include observations and measurements of a variety of types	0.72	0.66 ^b
			Students do work to transform raw data into evidence. They analyze data to produce graphs, models, and other data representations that illustrate trends and patterns within and across data	0.93	0.89 ^b
			Students reflect upon and verify data collection processes, accuracy of data, and transformation of evidence from data	0.68	0.63 ^b
Formulating evidence-based explanations	0.88	0.79 ^b	Students formulate explanations for causes of effect or establish relationships based on empirical evidence	0.85	0.80 ^b
			The formulated explanation fully answers their investigation question	0.77	0.72 ^b
			The formulated explanation illustrates learning	0.91	0.86 ^b
			The formulated explanation builds on their existing knowledge	0.90	0.86 ^b

a. Significant at the 0.005 level.

b. Significant below the 0.001 level.

repeated measures and data nesting using a multi-level mixed model analysis (Littell, Milliken, Stroup, Wolfinger, & Schabengerger, 2006). Our use of a mixed model analysis was not to examine the effects within the hierarchical structure of the data but rather to reduce the noise due to the nested data so we could determine if and where effects existed. The repeated measures were the three lessons (both enacted and planned) per teacher so our analysis grouped the per teacher enactments separate from the per teacher lesson plans.

We performed the mixed model analyses in SAS using a single-factor and double-factor repeated-measures mixed model ANOVA. We ran single-factor measures ANOVA to examine the difference in the submeasures for each individual feature (*giving priority to evidence* and *formulating evidence-based explanations*). In each of these ANOVA's the

dependent variable was the P-SIO scores for the teachers' lessons (enacted or planned) and the independent variable was the four submeasures within each individual feature. Our single-factor repeated-measures mixed model ANOVA formula is $Y_{ij} = \pi_{oj} + e_{ij}$, where Y_{ij} is the average of the three lessons (j) for each teacher (i); π_{oj} are the individual lessons per teacher (enacted or planned); and e_{ij} is the error in Y (Littell et al., 2006). Paired sample t tests were used to make *post hoc* comparisons between the submeasures to determine the source of the statistical significance.

We ran two double-factor repeated measures ANOVA addressing the differences between the planned (lesson plans) and enacted (video recorded) lessons using P-SOP scores for the feature *giving priority to evidence* and for *formulating evidence-based explanations*. In each of these ANOVA's, the dependent variable was the P-SOP scores for the lessons. The first independent variable is planned lessons versus enacted lessons and the second independent variable is the four submeasures that composite to form the measure feature score. Our double-factor repeated measures mixed model ANOVA formula is $\pi_{oj} = \beta_{oo} + \lambda_{oj} + r_{oj}$, where π_{oj} denote observations for individual o (teacher lessons) in group j (enacted or planned); β_{oo} is the mean for the enacted or planned lessons; and λ_{oj} is the effects for treatment (planned or enacted). The errors are represented both with variance and covariance as r_{oj} (Littell et al., 2006). If we found an interaction effect for the feature measure, we then ran a simple effects test by submeasure for planned and enacted lessons.

Embedded Multiple-Case Study

Case Study Participants

The four case study teachers were purposefully sampled (Creswell & Plano Clark, 2011) from the eight case study teachers participating in the multi-year professional development program. These four teachers were selected because all had been using their science curriculum materials for at least 2 years. The four case study teachers chosen for this study—two 3rd-grade and two 4th-grade—had the most extensive elementary teaching experience (range: 16–34 years; \bar{x} = 24.2 years) out of the case study participants and all held graduate degrees. The case study teacher demographics are outlined in **Table 2**.

Data Sources and Collection

In addition to live observations of these four teachers' enacted lessons, we also conducted a series of in-depth, lesson-specific, semi-structured (Patton, 2001) interviews with each case study teacher. The first occurred

Table 2. Case study teacher profiles

	<i>Grace</i>	<i>Helen</i>	<i>Emily</i>	<i>Danielle</i>
Additional education	MS Reading Literacy	MS Geoscience	MS Teaching and Leadership	MS Education
Classroom	Self-contained	Science/math	Self-contained	Self-contained
Years teaching	34	20	16	28
School	Eastwood	Southwood	Eastwood	Westwood
Free/reduced lunch	59.3%	73.8%	59.3%	32.3%
Grade level	3rd grade	4th grade	3rd grade	4th grade
Observation 1	Seeds	Craters	Seeds	Properties of substances
Observation 2	Seeds	Landforms	Seeds	Properties of substances
Observation 3	Seeds	Circuits	Seeds	Circuits

shortly before they enacted their lesson and emphasized their thoughts on their lesson plan, if and how their planned lessons engaged students in scientific practices, and any specific modifications they planned to make to their lesson. These interview questions included, in the preinterview, asking the teachers how their “original curriculum engaged students in collecting and transforming data” and if they thought their “original curriculum materials engaged students in formulating an explanation from their data and evidence.” The post-enactment interviews focused on how they interpreted their lesson enactment, the extent to which they considered scientific explanations emphasized in the lesson, and any modifications they made before, during, and/or after the lesson enactment. Questions included how they thought the lesson enactment “engaged students in collecting and transforming data” and how did they think the lesson enactment “engaged students in formulating and explanation from their data and evidence.” Both interview protocols were explicitly aligned with the theoretical framework underlying the study. Even though there was no difference in the questions we asked these teachers, we had a range of interview lengths that lasted on average for 19 minutes (range: 9:11–26:48 minutes). A variation in time was due to the teacher’s availability and willingness to elaborate on their responses. All interviews were audio-recorded and transcribed verbatim. Additionally, each case study teacher participated in an in-depth semistructured interview (Patton, 2001), once at the beginning and once at the end of the school year. On average, these interviews were 28 minutes in length (range: 17:07–45:22 minutes). Unlike the lesson-specific interviews, these two interviews focused on the teachers’ current understanding of scientific practices in the elementary classroom, particularly explanation construction.

All interviews were conducted by one of the authors either in person at a location of the teacher's choosing or on the phone.

Each teacher also submitted a number of other artifacts as data. For each of the three lessons, the case study teachers completed a lesson planning artifact (see Forbes & Davis, 2010) in which they evaluated how well the original lesson afforded students opportunities to *give priority to evidence* and *formulate evidence-based explanations*, suggested lesson adaptations to address these weaknesses, and justified their curriculum design decisions. Finally, each case study teacher completed a lesson plan evaluation of an example elementary magnets lesson that had been adapted to meet all five features of inquiry (Zangori et al., 2012). The evaluation asked the teachers to assess how inquiry-oriented they found the lesson and in what ways it did (or did not) meet each of the five features of inquiry. They submitted these electronically to the project team early in the year prior to planning and teaching their three science lessons. A separate interview was conducted with the case study teachers to discuss their lesson plan evaluations. During these interviews, teachers were asked to elaborate upon some of their critiques of the magnets lesson.

Data Analysis

The qualitative portion of this study is a holistic, multiple-case design with an emphasis on cross-case analysis (Yin, 2009). All 178 transcriptions and written artifacts from the four teachers were imported into *Atlas.ti* and coded by two of the authors for all five essential features of inquiry (NRC, 2000). Since our interviews focused on all five essential features of inquiry, they also served as our general coding scheme. We identified these five features of inquiry (question, evidence, explanation, alternate explanation, communicate/justify) as our five codes for the first coding level (Miles & Huberman, 1994). The lead author and second author jointly coded 20% of the text data sources. Inter-rater reliability among the texts averaged at 85% and, after discussion among the raters, a 100% agreement was reached. Source triangulation occurred through multiple data sources (e.g., interviews, teacher artifacts, and lesson observations) used in the analysis. The interviews, and lesson plans (both enacted and planned) were grouped by teacher by lesson for qualitative analysis to provide a more complete picture of the teachers' ideas about the inquiry features.

Once all of the data were coded, we performed code queries for codes *giving priority to evidence* and *formulating evidence-based explanations*. The purpose of this data reduction phase was to isolate data relating

Table 3. Case study coding scheme

<i>Coding Level 1</i>	<i>Coding Level 2</i>	<i>Emergent Themes</i>
Giving Priority to Evidence	Engagement with phenomena	Essential elements of scientific explanations
	Data collection	
	Data analysis	Their definition of an explanation (This submeasure was not discussed by the case study teachers)
	Data reflection/verification	
Explanation	Based on evidence	Beyond student developmental capabilities
	Answers investigation question	
	New understanding	

specifically to these two features of inquiry which are the focus of the study. Each of these two sets of queried data were then subjected to a second round of coding by the lead author using each feature's respective submeasure item descriptions as codes as outlined in **Table 3**. The objective of the second round of coding was to identify patterns (Miles & Huberman, 1994; Yin, 2009) within and among the case study teachers that illuminated how their pedagogical reasoning about *giving priority to evidence* and *formulating evidence-based explanations* informed the ways in which they engage students in these features of inquiry in the classroom. Data were queried and analyzed for each teacher and used to produce individual case summaries. Once we established the themes per teacher, we then performed pattern matching across the four case study teachers to create a cross-case synthesis regarding the pedagogical reasoning the four case study teachers exhibited regarding scientific explanations (Yin, 2009).

Results

Results from quantitative analysis of teachers' *planned* science lessons show that the science curriculum materials the participants used predominately emphasized *giving priority to evidence*, particularly the submeasures *engagement with the phenomena* and *data collection*, but rarely emphasized students' *formulation of evidence-based explanations*. Teachers' lesson enactments closely aligned with planned lessons so that students' engagement in *giving priority to evidence* and *formulation of evidence-based explanations* in the classroom closely mirrored opportunities

Table 4. Descriptive statistics for teachers' planned science lessons ($n = 121$)

Features and Sub-Measures	\bar{x}	Std. Dev.	Score Range		Frequency of Zero Score (%)
			Low	High ^a	
Giving priority to evidence	5.94	1.78	0	11	2.5
Engagement w/phenomena	2.34	0.76	0	3	4.1
Data collection	1.92	0.89	0	3	10.7
Data analysis	1.32	1.11	0	3	35.5
Data reflection and verification	0.30	0.63	0	2	79.3
Formulating evidence-based explanations	2.30	1.86	0	10	38.8
Constructing explanations based on evidence	1.08	1.08	0	3	41.3
Answers investigation question	0.50	0.85	0	3	68.6
New understanding	0.38	0.73	0	3	76.0
Existing knowledge	0.37	0.66	0	3	71.9

a. These are the lowest and highest scores for the planned lessons. The highest scoring lesson was an 11 out of a possible 12 points.

afforded them in the curriculum materials. Qualitative findings from the four case study teachers suggest two primary reasons for the emphasis on *giving priority to evidence* over students' *formulation of evidence-based explanations* in the enacted science lessons. First, the teachers viewed scientific explanations as only embodying the first two facets of *giving priority to evidence*—*engagement with the phenomena* and *data collection*—without considering *data analysis* as a necessary facet of scientific explanation construction. Second, they expressed concerns that their students would be able to engage in scientific explanation construction beyond *data analysis*. Both factors limited the opportunities afforded students to articulate mechanisms for observed phenomena beyond those already included in the curriculum materials.

Giving Priority to Evidence and Formulating Evidence-Based Explanations in Planned and Enacted Science Lessons

To address Research Question #1, we investigated two elements. First, we examined the degree to which the teachers' planned and enacted lessons emphasized the two features of inquiry foregrounded in this study. We found that overwhelmingly, the lesson plans emphasized students' *giving priority to evidence* (Table 4, Figure 1; $\bar{x} = 5.94$; $\sigma = 1.78$) much more so than students' *formulation of evidence-based explanations* (Table 4, Figure 1; $\bar{x} = 2.30$; $\sigma = 1.86$). We also examined the teachers' lesson enactments and observed the same trend (Table 5, Figure 1). The

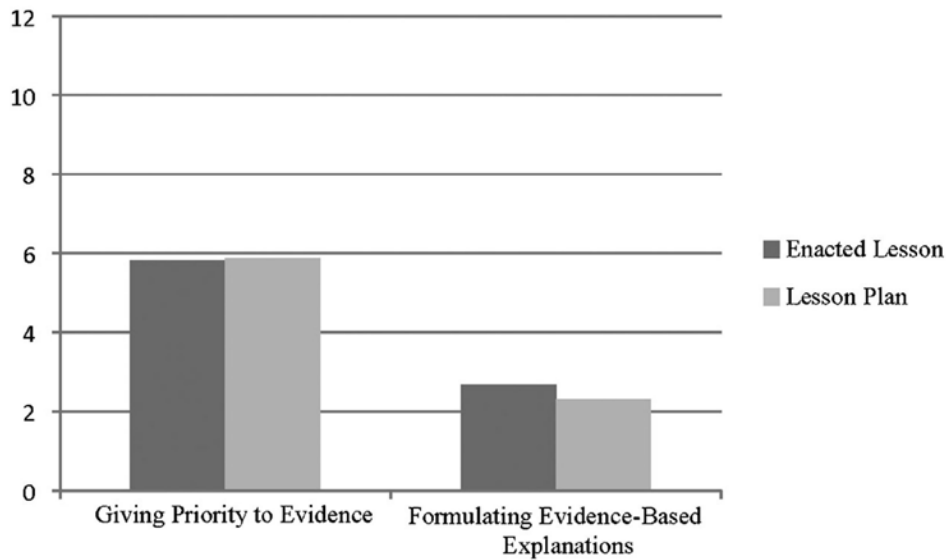


Figure 1. Mean scores for giving priority to evidence and formulating evidence-based explanations in teachers' planned and enacted science lessons.

Table 5. Descriptive statistics for teacher's enacted science lessons ($n = 121$)

Features and Sub-Measures	\bar{x}	Std. Dev.	Score Range		Frequency of Zero Score (%)
			Low	High ^a	
Giving priority to evidence	5.83	2.21	0	11	0.8
Engagement w/phenomena	2.28	0.55	0	3	1.6
Data collection	1.93	0.79	0	3	8.1
Data analysis	1.14	0.92	0	3	26.6
Data reflection and verification	0.52	0.72	0	3	58.1
Formulating evidence-based explanations	2.68	2.58	0	10.5	21
Constructing explanations based on evidence	1.06	0.89	0	3	27.4
Answers investigation question	0.58	0.88	0	3	60.5
New understanding	0.49	0.70	0	3	57.3
Existing knowledge	0.56	0.71	0	3	46.8

a. These are the lowest and highest scores for the enacted lessons. The highest scoring lesson was an 11 out of a possible 12 points.

measured differences between the presence of *giving priority to evidence* and *formulation of evidence-based explanations* is statistically significant for both lesson plans, $t(39) = 12.37$; $d = 1.99$; $p < 0.0001$, and enacted lessons, $t(39) = 4.23$; $d = 2.07$; $p < 0.0001$.

Planned Lessons

We examined the extent to which each of the four submeasures for *giving priority to evidence* and *formulation of evidence-based explanations* was emphasized in the teachers' planned lessons. First, we examined the differences in the submeasure means for *giving priority to evidence* and found a significant effect for the prevalence of the four submeasures within the teachers' planned lessons, $F(3, 351) = 178.19, p < 0.0001$. Paired sample t tests were used to make *post hoc* comparisons between the mean submeasure scores. We found that the differences between the mean of each submeasure were statistically significant. This significant effect can be seen in Figure 2 and Table 6.

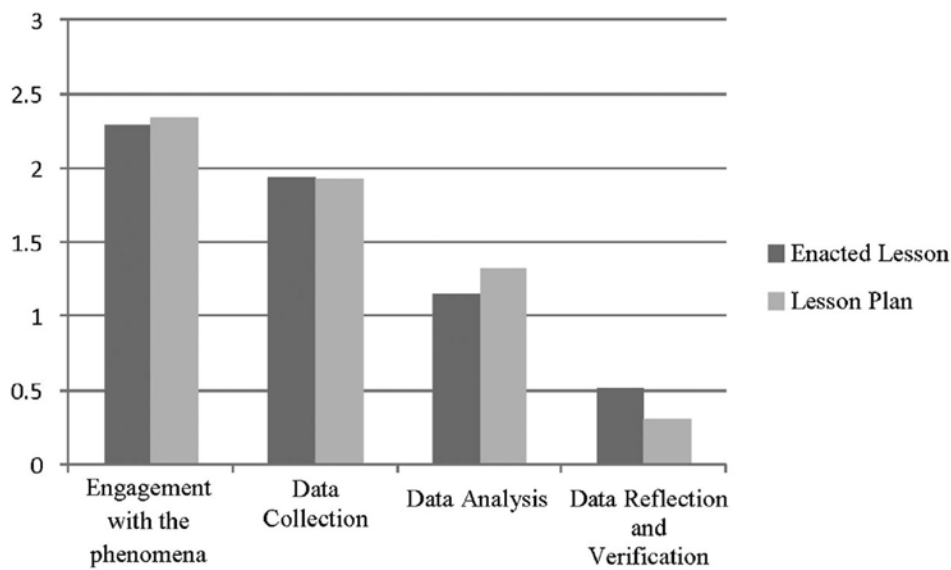


Figure 2. Mean scores for giving priority to evidence submeasures in teachers' planned and enacted science lessons.

Table 6. Statistical comparisons between submeasures for giving priority to evidence

Submeasure Comparison	Lesson Plans			Enacted Lessons		
	t	p	Cohen's d	t	p	Cohen's d
Engagement with phenomenon and data collection	4.46	<0.0001	0.51	4.75	<0.0001	0.514
Engagement with phenomenon and data analysis	10.87	<0.0001	1.07	15.13	<0.0001	1.504
Engagement with phenomenon and data reflection and verification	21.74	<0.0001	2.922	23.34	<0.0001	2.75
Data collection and data analysis	6.42	<0.0001	0.59	10.37	<0.0001	0.54
Data collection and data reflection and verification	17.29	<0.0001	2.10	18.59	<0.0001	1.87
Data analysis and data reflection and verification	10.87	<0.0001	1.130	8.21	<0.0001	0.75

To determine the frequency that each submeasure appeared within the planned lessons, we investigated the percentage of zero scores for individual submeasures. A score of zero on the PSOP indicates that the submeasure was not present in the lesson. We subtracted the percentage of zeros within the submeasure from 100% to determine the frequency that the particular submeasure appeared overall in the total scored planned lessons. We found that the evidence submeasure *engagement with phenomena* was by far the most prevalent in 96% of the total planned lessons. The next most frequent submeasure was *data collection*, which was observed in 89% of the planned lessons. The remaining two evidence submeasures were observed less frequently at 64.5% for *data analysis* and 20.7% for *data reflection/verification*.

Second, we examined the differences in the submeasure means for students' *formulation of evidence-based explanations* and also found a significant effect, $F(3, 351) = 39.11, p < 0.0001$, overall among differences in means. Paired sample *t* tests were used to make *post hoc* comparisons between each of the mean submeasure scores. Per our *t* tests analysis, the lesson plans emphasized *constructing explanations based on evidence* at a statistically significant level ($p < 0.0001$) over the other submeasures for students' *formulation of evidence-based explanations* (**Table 7**). This significant effect is visible in the mean scores on Figure 3 where the mean of *constructing explanations based on evidence* is approximately twice that of the other submeasures and its frequency was the most prevalent, observed in 59% of the enacted lessons. The remaining submeasures were observed much less frequently (range: 24–31%).

Table 7. Statistical comparisons between submeasures for formulating evidence-based explanations

Submeasure Comparison	Lesson Plan			Enacted Lessons		
	<i>t</i>	<i>p</i>	Cohen's <i>d</i>	<i>t</i>	<i>p</i>	Cohen's <i>d</i>
Constructing explanations based on evidence and Answers investigation question	7.54	<0.0001	0.60	7.39	<0.0001	0.54
Constructing explanations based and New understanding	9.43	<0.0001	0.76	8.57	<0.0001	0.71
Constructing explanations based and Existing knowledge	9.10	<0.0001	0.79	7.45	<0.0001	0.62
Answers investigation question and New understanding	1.89	<0.0001	0.15	1.18	<0.0001	0.11
Answers investigation question and Existing knowledge	1.55	<0.0001	0.17	0.06	<0.0001	0.03
New understanding and Existing knowledge	-0.33	<0.0001	0.014	-1.12	<0.0001	-0.09

Enacted Lessons

Next we examined the extent to which each of the four submeasures for *giving priority to evidence* and *formulation of evidence-based explanations* was emphasized in the teachers' enacted lessons. We found a significant effect for *giving priority to evidence* within the enacted lessons, $F(3, 360) = 219.49, p < 0.0001$. We again used paired samples *t* tests to make *post hoc* comparisons between the submeasure scores. The differences between the means of each submeasure score were statistically significant (*t*-statistic range: 8.21–23.34; $p < 0.0001$; see Table 6 and Figure 2). The enacted lessons reflected slightly different frequencies in submeasures *data analysis* and *data reflection/verification* than did the planned lessons. *Data analysis* occurred more frequently in the enacted lessons than in planned lessons; however, as we discuss in the comparison between planned versus enacted lessons below, the frequency of data analysis here was not statistically significant.

We also found a significant effect, $F(3, 360) = 31.06, p < 0.0001$, among the means of the submeasure scores for *formulation of evidence-based explanations*. The *post hoc* comparisons among the submeasure means using paired sample *t* tests found that, just as with the planned lessons, the enacted lessons also emphasized *constructing explanations based on evidence* at a statistically significant level ($p < 0.0001$; see Table 7 and Figure 3) over any of the other submeasures for *formulation of evidence-based explanations*. Also as with the planned lessons, this significant effect can be seen in the mean scores on Figure 3 where the

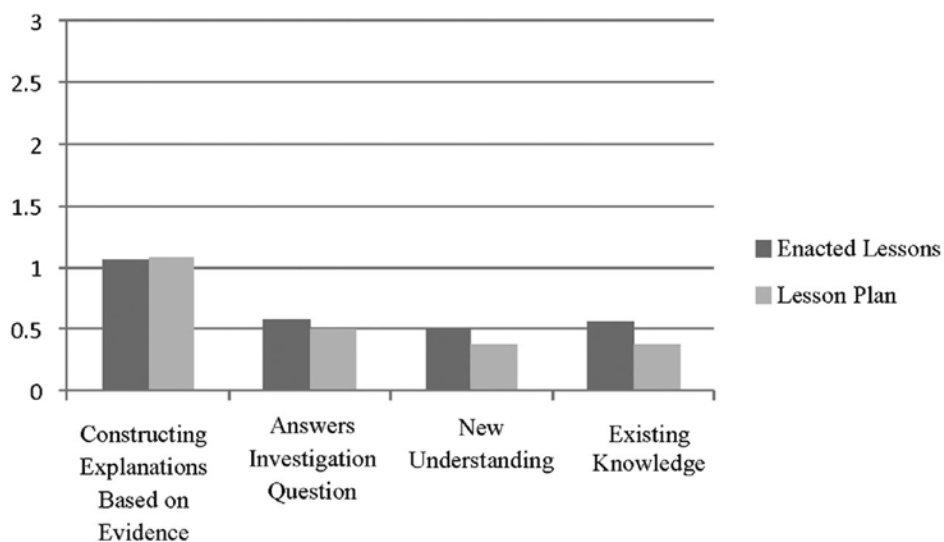


Figure 3. Mean scores for formulating evidence-based explanations submeasures in teachers' planned and enacted science lessons.

enactment mean of the submeasure *constructing explanations based on evidence* is twice that of the other enacted submeasures. According to the calculated frequency, this submeasure was observed in 72% of the enacted lessons. The remainder of the submeasures for *formulation of evidence-based explanations* were observed much less frequently (range: 39.5–53.2%). The frequency of these submeasures is slightly higher than those observed in the planned lessons.

Planned Versus Enacted Lessons

Next, we examined if the enactment submeasure means for *giving priority to evidence* and *formulation of evidence-based explanations* differed significantly from the curriculum materials. While we did not find a significant main effect for planned and enacted lessons for *giving priority to evidence*, $F(1, 828) = 0.01$, $p = 0.918$, we did find a main effect for differences among the presence of the submeasures in planned versus enacted lessons, $F(3, 828) = 366.4$, $p < 0.0001$, and observed an overall significant interaction effect, $F(3, 828) = 369.4$, $p = 0.0117$, between the enacted and planned lessons. In other words, while there was no difference in the overall presence of *giving priority to evidence*, there was a significant difference among and interactions between one or more of the submeasures in teachers' planned and enacted lessons.

As can be seen by the submeasure means in Figure 2, the overall significant interaction between the submeasures was attributed to the mean differences for submeasures *data analysis* and *data reflection/verification*. The mean for *data analysis* in lesson plans is higher than the enacted lessons, meaning that *data analysis* was observed more frequently in the lesson enactments than in the lesson plans, but as found in the one-way ANOVA (i.e., simple effects test), this difference was not statistically significant. The means for *data reflection/verification* had an opposite occurrence—data reflection/verification was observed less in the enacted lessons than in the planned lessons. This difference between the planned and enacted means for the submeasure *data reflection/verification*, $F(3, 828) = 6.89$, $p = 0.0088$, $d = 0.32$, was statistically significant. However, and importantly, the difference between the planned and enacted means for *data reflection/verification* was not large (planned: $\bar{x} = 0.37$; enacted: $\bar{x} = 0.57$) and both scored very low for the presence of the submeasure. This suggests that this feature was so minimally present in the planned lessons that even very slight modifications by the teachers resulted in a significant finding.

For *formulation of evidence-based explanations*, we found a significant main effect in the difference between the means of the planned and

enacted lessons $F(1, 828) = 7.15$; $p = 0.0076$, and in the differences between the means of the submeasures, $F(3, 828) = 51.89$; $p < 0.0001$. No significant main effect interaction between the submeasures and type of lesson (planned or enacted) was observed. Paired t tests were used to make *post hoc* comparisons between the explanation submeasures. The submeasure *constructing explanations based on evidence* occurred at a higher level in both the planned and enacted lessons than any of the other three submeasures ($p < 0.0001$; see Table 7). Differences among the three other submeasures were not statistically significant.

Elementary Teachers' Pedagogical Reasoning About Giving Priority to Evidence and Formulating Evidence-Based Explanations

To address Research Question #2, we examined how teachers' pedagogical reasoning explains their emphasis on particular aspects of *giving priority to evidence* and *formulating evidence-based explanations* in their use of science curriculum materials in the classroom. Findings from the multiple-case study of Grace, Helen, Emily, and Danielle suggest two main themes. First, for *giving priority to evidence*, the teachers emphasized *engagement with the phenomena* and *data collection*, but typically not *data analysis*, as the crucial and necessary facets of scientific explanation construction. Second, the teachers did not emphasize mechanism as part of *formulation of evidence-based explanations*, instead focusing on *data analysis* as the defining element, in part because they viewed this practice as beyond the developmental capabilities of their students.

Giving Priority to Evidence

All four teachers identified *engagement with phenomena* and *data collection* as essential elements of giving priority to evidence as part of scientific explanation construction. Throughout the year, the teachers emphasized that students need to spend the bulk of their time during a science investigation *engaging with the phenomena* and in *data collection* so they had opportunities to "act like scientists" and be "more scientific" (Emily, P36:138). The teachers focused on *engagement with the phenomena* because, as they described, it was taking part in "hands on," materials-rich science activities that defined the science experience for students. When students were observing, recording their observations, and working in small groups, then they were engaged in scientific activities. For example, when asked about the importance of students' working with data and evidence, Danielle stated:

I think it's very important they record their results. . . I tell them all the time that's what scientists do and. . . that's why it's important because I expect them to go on a become scientists (P41:50).

As Danielle expressed, it is *engagement with phenomena* and *data collection* where students are provided opportunities to do “what scientists do.”

However, none of the case study teachers included in their discussions how they provide opportunities for students to organize, analyze, or represent their data in their definitions of ways in which students “act like scientists” When we specifically asked about *data analysis* in their formal interviews, Helen and Emily suggested that it was an important piece of a classroom investigation but were unable to express specifically in what ways they could provide opportunities for data analysis during their lessons. Instead, Helen and Emily identified that analyzing data provided students with the “why” of the lesson—which they equated as the “correct” answer. In their definition of *data analysis*, students should “look at it, be able to read it and understand it and transfer that information into an answer in a question” (Emily, P31:84) because this was when students would be able to understand the science content. As Emily stated, data analysis is important because it was when students discovered whether their results were “correct or incorrect and then *why* they were correct or incorrect” (Emphasis added, Emily, P37:149) When we inquired how the students were engaged in examining “why” they were correct during data analysis, Emily and Helen consistently noted that during *data analysis* students had the opportunity to examine what went wrong with their investigations and begin to look at variables such as how they setup and conducted the investigation. As Helen stated, “the variable piece is huge” (Helen, P29:127) and analysis of data makes the lesson “more authentic” (Helen, P23:104) because it is when students learn the science content. And as both Helen and Emily stressed, learning the science content was the most important part of the lesson.

Danielle did include organizing and analyzing data in her definition of “acting like scientists” because she found that if she did not have her students doing something more with their data then “there's just no reason for them to be doing it, it's just busy work” (P41:63), recognizing the importance of making sense of the data in order for it to be meaningful. However Danielle, like Emily and Helen, also focused on the importance of what the results of the investigation *should be*. Danielle was

very concerned with students getting the correct answer from their science investigations. As she expressed:

How can they learn the content unless I say, ‘This is what you should—this is what should have happened on your experiment,’ and then we can talk about why it didn’t. And I think if you’re teaching content, they do have to know the right answer. They need it for the test (P39:232).

As Danielle’s statement illustrates, the teachers’ focus on the evidence feature involved more than just “acting like a scientists” and extended to “getting the right answer.” While Danielle was the only case study teacher that included the importance of data analysis in supporting students to get the right answer, she aligned with Helen and Emily’s concerns that if the collected data did not match what the answer should be (e.g., seeds in water sprout), then they have to help their students determine where the experimental error occurred and stress the “correct” result that should occur if there was no error. Again, while experimental error is an important component of scientific research, the teachers did not include the mechanism for *why* the “right” answer would occur (e.g., seeds are living things and require water to grow).

While Grace also heavily embodied the idea that *engagement with the phenomena* and *data collection* was “doing science,” she differed from the other teachers in that she did not include *data analysis* in her definition of the evidence feature at any time during the study. In her words student engagement in the engagement and data collection submeasures is “put[ing] the ownership on the kids ... to find their own evidence” (P51:188) so as to mimic real world science. Grace, even more than the other three teachers, embodied the perspective that when students are active with the phenomena and writing down measurements, then they are “doing science” and mimicking scientists—or as she stated mimicking “real world” (P49:112) science investigations. Her discussions throughout our interviews never included *data analysis* as part of “doing science,” even when we asked her specifically how opportunities for *data analysis* might look during her lessons. When Grace was asked directly about *data analysis*, her responses emphasized the continuing need to engage students in discussions about their observations so that they could better engage in the practices of science.

Observations of the teachers’ enacted lessons indicated that, in some instances, the teachers’ espoused pedagogical reasoning did not align with their enactments. First, as already discussed, Emily and Helen’s

discussions did not include a focus on *data analysis* in order for students to learn the science content. Across all three of Emily and Helen's lesson enactments, only one lesson for each of these teachers included data analysis activities. However, the difference between Helen and Emily is that while Emily closely followed her curriculum materials, which included some elements of data analysis in one lesson, Helen heavily modified her original materials to include data analysis where the original lesson plan did not. Helen was an exceptional case in that she was the only case study teacher that modified her original curriculum materials. She re-wrote her materials so extensively that in her lesson enactments, the lessons she presented only had the subject matter in common with the original curriculum. In her third lesson, for example, a kit-based magnets unit that she referenced in her initial interview, she substantially modified opportunities for student engagement, data collection, and data analysis by breaking the student activities into six workstations that she described as "discovery boxes" and included a different investigation question about magnets at each one. She rotated small student groups through each work station where they had four minutes to gather data, do an analysis, and determine *what* happened and *how* it happened to answer each individual investigation question. She asked the students to both discuss and write in their notebooks a description of *what* they observed and *how* they thought it might have occurred. After all students had engaged in all workstations, she brought students to the carpet in the center of the room asking them "What are we going to do with the information we collected?" (Helen, 2/b:14:51) and supported her students to understand the cause and effect.

In contrast, Emily made no modifications to her original curriculum materials and enacted her lessons just as they appeared which, for Emily did not include any additional data analysis than what was only minimally present in the curriculum materials. Danielle also enacted her curriculum without modification, but the curriculum materials for all of her enactments heavily stressed data collection, while Emily's did not. As a result, we observed data analysis occur in each of Danielle's lesson (**Figure 4**). For example, in Danielle's second enactment about properties of substances she mentioned prior to the lesson that it was very effective at developing sense-making skills which are also included on the mandated district science test. As she stated "there's no way I [can] prepare [students'] ... unless I directly teach to the test" (Danielle, P39:036), which she interpreted as closely following her curriculum materials. This lesson involved using three different dishwashing detergent-water mixtures for students to evaluate by measuring the size of the bubbles they

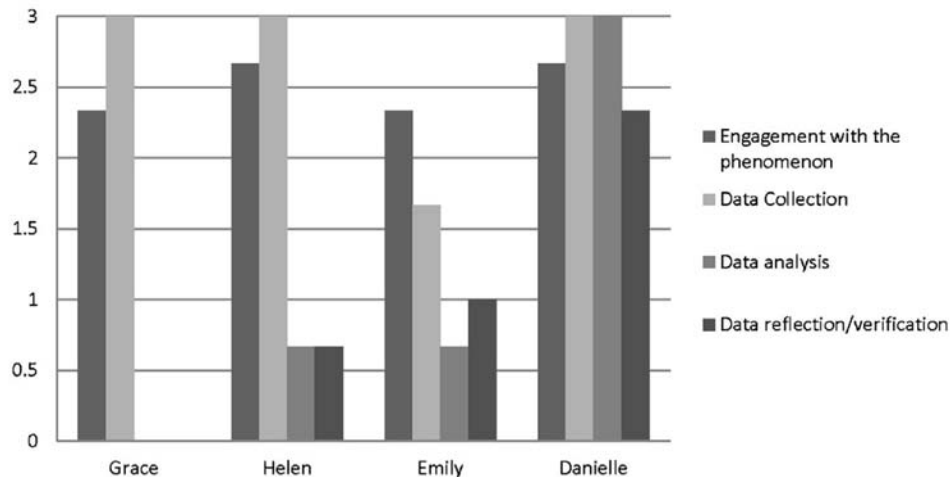


Figure 4. Giving priority to evidence submeasure mean scores for each case study teacher.

can make on a flat surface using a straw. The students blow bubbles three times with each detergent and then measure the size of the bubbles using the “shadow” that remains after the bubble pops. Danielle provided a data sheet that was included with her second lesson where students recorded their data (three trials per experiment) and then averaged their trials. She asked each student group how they should interpret their data then encouraged them to explain what their data means to each other. She provided opportunities for her students to verbally describe, to her and to classmates, *what* happened in their investigation—to link cause and effect—thereby effectively providing opportunities for her students to engage in *data analysis*. Further, she also strongly engaged her students in data reflection and verification—more so than we observed in any other lesson included in this study—because, as she described, it was important that her students focus on what the results should be, and if they did not get those results, then it was important they determine where their experimental error occurred.

Grace’s pedagogical reasoning about the importance of *engagement with phenomena* and *data collection* was strongly evident in her lesson enactments (Figure 4). For example, Grace taught a series of three lessons that afforded students opportunities to investigate the relationship between seeds, plants, and food. In her first lesson, students described, dissected, and counted the number of seeds in a bean pod. In her second lesson, students were to dissect fruits and locate, count (or estimate) seed number, and sort the seeds by properties. Grace made few modifications to the original versions of the lessons and closely followed each

step of her curriculum materials, except for when the curriculum materials called for *data analysis*. For example, in Grace's second lesson, she followed the lesson plan closely which emphasized students' *engagement with the phenomena* and, to a slightly lesser extent, *data collection*, just as it appeared in the original curriculum materials. However, she did not enact the latter parts of the curriculum materials where the students were instructed to begin analyzing their data to look for patterns and relationships among the different seeds from different fruits. After the lesson, when asked if she would engage students in this part of the lesson later, Grace noted that her students "are not anywhere near ready to analyze that data" (Grace, P52:100) because, as she described, they had not yet collected enough data to analyze.

Formulating Evidence-Based Explanations

Each of the case study teachers illustrated varied ideas about and understandings of *formulating evidence-based explanations* as a scientific practice, but most frequently their articulated definitions of explanation construction aligned with *data analysis*. They identified that when student "explaining" occurred, it involved students interpreting graphs or other methods of data analysis to summarize patterns or trends. In other words, when students were explicating their data, the teachers considered they were engaged in explanation construction. When asked how important it was to provide opportunities for students to formulate evidence-based explanations, they struggled to articulate what was meant by "explanation." For example, Grace and Emily responded that "explanations" were the point in the lesson were students discussed their data as "their own discoveries" (Grace, P49:183) and if the students were unable to discuss their data "then they probably haven't understood the information" (Emily, P31:86). Danielle also responded in alignment with Grace and Emily that without a discussion of the data, then "data has . . . no meaning" (Danielle, P41:61). In each of these discussions on the importance of formulating evidence-based explanations, the teachers conflated an explication of data with an explanation and focused instead on students describing their data and evidence in order to determine what happened—a cause and effect. However, in none of these three teacher's discussions did they include an examination of the underlying, unobservable mechanism for the cause and effect and how they might provide opportunities for their students to engage in scientific sense making.

Interestingly, Helen was the only teacher that extended her definition of explanations past explication to emphasize mechanisms for cause and effect that might help students understand how or why

phenomena occur. In response to our question about the importance of including opportunities for students to *formulate evidence-based explanations*, she stated:

Students need to think about the possible reasons *why* something is the way it is, or *how* they got to that answer. . . explanations should be based on a combination of prior knowledge and what they found. The reason should be directly related to the purpose set by the inquiry. (Emphasis added; Helen, P25:36)

Helen went on in this discussion to specifically identify the necessity of the mechanism within scientific explanations, using a magnet lesson that she planned to do in the upcoming school year as her example. She discussed how she hoped to connect student evidence on magnets to the underlying science concept of positive and negative poles (P22:128). Helen spoke of all of the components of explanation: cause and effect that she expressed as “what they found [evidence]” and mechanism that she expresses as “why something is the way it is or how they got that answer.” Helen was the only teacher that emphasized all of the components that define a scientific explanation, though she only expressed explanations in this manner during her initial interview. In her following interviews, her perspectives on explanation aligned with those of the other three teachers that highlighted explanations as explications of *data analysis*.

We did find, however, that despite the conflation of explication with explanation, all four teachers engaged their students in construction of scientific explanations and opportunities for sense making to varying degrees in the classroom (**Figure 5**). Emily, Grace and Danielle each in two lesson observations and Helen in one lesson observation. Helen even made changes to curriculum materials in the one lesson that emphasized this feature to better engage students in *formulating evidence-based explanations*. As we discussed earlier, she modified her third lesson substantially. This lesson was a kit-based magnets unit in which she altered the lesson to involve work-stations in which she rotated small student groups through each work station. Each of five work stations had different investigation questions related to magnets with the sixth and final work station asking the cumulative question of “How do magnets work?” In order for students to answer this question, they must examine the cause and effect they had observed at the other five work stations to propose the mechanism. Helen did not engage her students during their visits to work stations but instructed them to ask their group members if they had questions. The group questions predominately focused on

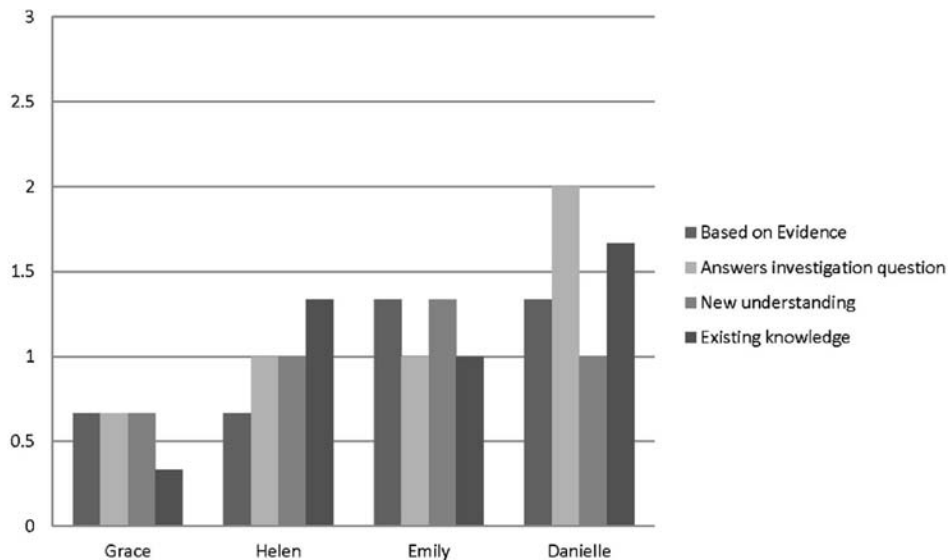


Figure 5. Formulating evidence-based explanation submeasure mean scores for each case study teacher.

showing the other students what the magnets did and trying to get the magnets to do different things.

After all students visited all work stations, Helen invited them to the carpet in the middle of her room to discuss their findings. Rather than inviting her students to link cause, effect and mechanism, she focused student discussions on explicating their observations and did not move them to propose mechanisms. While this lesson had the potential for students to link cause and effect with mechanisms, particularly with her conception of pulling together all of their information for station six, the enactment did not score high on our rubric (Figure 5). This was because even though the potential was there for cause, effect and mechanism, we did not observe students engage in sense making about magnets in the discussion nor did we observe Helen attempt to support her students in linking cause, effect and mechanism. Instead the student discussions and her questions of her students focused on data explication.

As we saw with the evidence feature, Danielle followed her curriculum materials closely during her lessons. Her second lesson included some opportunities for students to construct explanations; however, her third lesson heavily focused on having students connect cause, effect, and mechanism. For lesson three she told her students, as well as us, that they would need to understand all the particulars of their experiments including an understanding of “why” because “it’s going to help you [the students] on the [district assessment] test!” (Danielle, 3b/7:00). This lesson

also scored the highest overall for the presence of evidence-based explanations of all lessons enacted by the case study teachers. Frequently, during this lesson, Danielle asked her students to use their evidence to tell her “why” the circuit worked (or did not work) to light the bulb. This focus on scientific sense making was, as she stated several times during her lesson enactment, due to her concern about how the students would do on the assessment. Throughout the lesson she provided continuous support and feedback to her students about the mechanisms that caused the bulbs to light (or not light) until the class all came to an agreed upon “why”—the mechanism for the cause and effect. This is the only lesson we observed where students were provided multiple opportunities to engage in scientific sense making.

However, after this lesson, even though Danielle had just heavily supported her students in cause, effect, and mechanism, she expressed that she preferred not to engage her students in *formulating evidence-based explanations* because she felt it was beyond the developmental capabilities of her students. When asked how well she thought her students were able to engage in scientific explanations throughout the school year generally and specifically in her third lesson, she referenced the difficulty she felt she had in an electricity unit, requiring students to build upon their data analysis to explain how they were able to connect wires to a battery in order to light a light bulb and the time it took for students to determine a mechanisms. Danielle stated:

I think this is very higher level. It's a whole lot of applying what they learned and justifying it. And I just think they're only fourth graders. What I want. . . I mean, I would be very happy if they could just showmen that they could light a light bulb. . . And if they could showmen that they could get two bulbs glowing brightly I would be happy with that because this is their first time ever with electricity. But to me, to ask them to take it to that higher level, I think that's difficult (P45:183).

All of the case study teachers expressed similar concerns about students engaging in formulating evidence-based explanations in the elementary classroom. Towards the end of the study, we asked each of the teachers specifically about having students make connections between cause and effect articulated through data analysis and mechanism-based explanations during their science lessons. All four of the case study teachers described this activity as too complicated for their students. For example, Helen suggested that her students were developmentally unprepared to

successfully engage in making these connections, stating “I was reminded how very immature they are ... a couple of them didn’t pick up what I wanted them to pick up. Even ... with ‘how do people use magnets?’ they got lost” (P29:152). Grace and Emily both also discussed the immaturity of their students in being able to make those connections within a science lesson. They reiterated how much scaffolding it took for students just to make “good observations” (Emily, P31:70) and were unsure if students would be able to go further than data explications. As such, the teachers’ assumptions about developmental limitations of students influenced their pedagogical reasoning about students engaging in formulating evidence-based explanations in the classroom.

Summary of Findings

Quantitative analysis of the teachers’ planned lessons show that the elementary science curriculum materials used by the teachers emphasized students *giving priority to evidence over formulation of evidence-based explanations*. This difference was predominately due to the emphasis on components of *giving priority to evidence: engagement with the phenomena and data collection*. Results from the quantitative analysis of the teachers’ enacted science lessons show that the teachers’ enacted lessons closely followed their lesson plans, though they did make some instructional adjustments to better emphasize *data analysis* as well as some elements of students’ *formulation of evidence-based explanations*.

The qualitative analysis of the four teachers—Helen, Emily, Grace, and Danielle—illustrates two primary reasons why *formulating evidence-based explanations* was emphasized less in the teachers’ enacted lessons than *giving priority to evidence*. First, the case study teachers highlighted the hands-on portion of their science lessons, which largely focused on *engagement with the phenomena and data collection*, as the means to get their students engaged and excited about science. Second, their conceptions of explanation construction revolved around a focus on describing cause and effect for phenomena, or *what* happened, rather than underlying mechanisms, or *how* and *why* the phenomena occurred. The teachers largely emphasized *data analysis* as explanation construction. All four of the teachers expressed concerns about the developmental abilities of their students to engage in *formulating evidence-based explanations*. As a result, though some of the teachers made minor modifications to their science lessons that better engaged students in some aspects of explanation construction, their enacted lessons largely mirrored those included in the science curriculum materials they used.

Synthesis and Discussion

Elementary science learning environments should foster early learners' sense making about the natural world (Duschl et al., 2007; NRC, 2000, 2012). Recent research has shown that when effectively scaffolded to do so, elementary students are capable of productively engaging in scientific practices to formulate explanations about natural phenomena (Cavagnetto et al., 2010; Hapgood et al., 2004; McNeill & Krajcik, 2008; McNeill, 2011; Metz, 2011; Ruiz-Primo et al., 2010; Samarapungavan et al., 2008). We have drawn on a sizable sample of enacted science lessons to show how teachers used science curriculum materials to afford students opportunities to *give priority to evidence* and *formulate evidence-based explanations*. Evidence from this study suggests that in the elementary classrooms we observed, students were afforded less-effective and more infrequent opportunities to *formulate evidence-based explanations* than engage with and *give priority to evidence*. Further, elementary teachers' enacted science instruction provided students similar opportunities to *give priority to evidence* and *formulate evidence-based explanations* as were afforded by the elementary science curriculum materials they used. In both planned and enacted science instruction, there was a clear emphasis on *giving priority to evidence*, which includes activities such as investigation, data collection, and data organization, over explanation construction. This finding corroborates prior research that evidence-based explanations tend to be underemphasized in science instruction in elementary and middle school (Forbes et al., 2013; Kesidou & Rosemann, 2002; Magnusson & Palincsar, 2005; McNeill & Krajcik, 2008; Minogue et al., 2010; Ruiz-Primo et al., 2010). Our study findings highlight commonly-observed limitations of elementary science learning environments and begin to shed light on some of the potential reasons for these observed trends.

As illustrated in the overall findings from this study, the elementary science curriculum materials focused instruction more heavily on students *giving priority to evidence*, or identifying cause and effect for natural phenomena, with far fewer opportunities for students to engage in scientific sense making through the formulation of evidence-based explanations that link cause and effect with underlying mechanisms. Science curriculum materials are important tools with which teachers design science learning environments. Recent reviews of middle school (Kesidou & Rosemann, 2002) and secondary science curriculum materials (Beyer, Delgado, Davis, & Krajcik, 2009) find that curriculum materials do not strongly emphasize sense making or provide educative supports

for teachers to understand either the importance of or in what ways to support sense-making activities in the classroom. While a comprehensive review of elementary curriculum materials has yet to be conducted (Kesidou & Rosemann, 2002), evidence from our other studies suggest that the de-emphasis on explanation-construction is part of a larger trend in which commercially produced, widely-used elementary science curriculum materials may afford students limited opportunities to engage in substantive scientific sense making about natural phenomena (Biggers et al., in press; Forbes & Davis, 2010; Forbes et al., 2013). This is an important trend of which many elementary educators are tacitly aware but for which there has thus far been little empirical evidence. It is particularly troubling given the extent to which elementary teachers, who are typically generalists with limited content knowledge, rely on curriculum materials to engage students in scientific practices (Duschl et al., 2007). As such, this study begins to document potential limitations of existing, widely-used elementary science curriculum materials.

Our findings also show that the elementary teachers in this study rarely adapted their science curriculum materials to better engage students in these crucial sense-making practices. We found very few of the teachers spontaneously and/or strategically modified their curriculum to either emphasize or deemphasize student sense making. While we did see both teacher planned and improvised modifications during the enactments, they were minor and did not significantly alter the P-SOP scores of the original lesson plans. As a result, the enacted science instruction that was observed in these teachers' classrooms was most heavily characterized by students' engaging with phenomena through hands-on activities (*giving priority to evidence*) with limited emphasis on the formulation of evidence-based explanations. These findings reinforce those of other studies that have similarly found elementary science instruction to emphasize engaging, hands-on experiences for students in lieu of sense making (Beyer & Davis, 2008; Biggers et al., in press; Forbes et al., 2013; Hardy et al., 2006; King, Shumow, & Lietz, 2001). While past research has explored preservice elementary teachers' use of science curriculum materials (e.g., Forbes & Davis, 2010) and in-service elementary teachers' use of curriculum materials for mathematics (e.g., Remillard, 2005), little research has been conducted to investigate how in-service elementary teachers use elementary science curriculum materials. Findings from this study begin to address this gap in the literature by illustrating the limited ways in which elementary teachers adapt elementary science curriculum materials to support students' formulation of evidence-based explanations as part of everyday professional practice.

What factors might help explain the limited degree to which elementary teachers in this study modified the science lessons they taught? As shown in the case study findings, the teachers articulated fundamental misconceptions about the nature of explanation construction as a central scientific practice in the elementary classroom. They tended to emphasize components of classroom inquiry that revolved around students' engagement with phenomena as the core component of inquiry, a finding that corroborates results of past research on experienced elementary teachers (King et al., 2001; Metz, 2009; Minogue et al., 2010), new elementary teachers (Avraamidou & Zembal-Saul, 2005; Beyer & Davis, 2008), and preservice elementary teachers (Forbes & Davis, 2010; Zangori & Forbes, 2013). Findings presented here extend this literature, however, by showing how the teachers' ideas about scientific explanations and the practice of explanation construction influenced the instructional decisions they made. Even when engaging students in investigation and data collection, the teachers often believed that they were engaging in instruction that supported their students' explanation construction about target phenomena. In particular, the teachers often equated students' data analysis to explanation construction, where students' organization and description of observational trends supplanted the need for students to reason and make inferences. Consistent with Braaten and Windschitl's (2011) perspective on explanation construction, this resulted in an emphasis in the enacted lessons on *what* happened (cause and effect) rather than *why* or *how* it happened (mechanism).

The teachers in this study also described students' developmental abilities as potential barriers to engaging them in explanation construction. This view is aligned with powerful and pervasive perspectives on student learning (Duschl et al., 2007; Metz, 1995) that underlie the design of some of the very science curriculum materials used by the teachers in this study. Yet there is an increasingly robust literature base that demonstrates elementary students are capable of engaging productively in scientific practices to make sense of natural phenomena (Cavagnetto et al., 2010; Hapgood et al., 2004; McNeill, 2011; Metz, 2011; Samarapungavan et al., 2008; Songer & Wenk Gotwals, 2012). As illustrated here, there may be a fundamental mismatch between, on the one hand, contemporary research findings and, on the other, the resources and practices that constitute day-to-day work in elementary science learning environments. This finding highlights the need for teachers to see concrete examples of scientific sense making in elementary settings to understand how students' engagement in scientific practices can be actively supported. This is particularly crucial since students' ability to formulate evidence-based

explanations does not develop spontaneously through hands-on engagement with the phenomena alone. Instead, it requires explicit scaffolding through teacher prompts and curricular support (Hapgood et al., 2004; Hardy et al., 2006; Kesidou & Rosemann, 2002). As such, it is crucially important that elementary teachers not only develop a robust understanding of how explanation construction builds from students' engagement with phenomena, but also that elementary students are capable of articulating mechanism-based explanations for phenomena they observe and document, as well as how they can be actively supported to do so.

Implications

These findings have important implications for curriculum developers and the preparation of teachers at both the preservice and in-service stages of their careers. First, elementary science curriculum materials need to provide experiences for students that include opportunities for students to formulate evidence-based explanations. An increasingly robust body of work in science education has begun to illustrate specific strategies and methods for effectively fostering elementary students' explanation construction (e.g., Hapgood et al., 2004; Hardy et al., 2006; McNeill, 2011; Metz, 2011; Samarapungavan et al., 2008; Songer & Wenk Gotwals, 2012). These include a curricular focus on core concepts that provide depth over breadth while making the domain-specific language of science explicit (e.g., scientific explanations versus everyday explanations). Effectively designed elementary science curriculum materials provide students opportunities to engage with phenomena and observe cause and affect relationships as part of a broader process of identifying, representing, and testing mechanisms that explain patterns and trends in natural phenomena. Students should be afforded opportunities to explicate evidence-based explanations about natural phenomena through a variety of representations, including writing, diagrams, physical models, and discourse. Transitioning between social meaning-making (involving students and teachers) and individual reflection while negotiating multi-modal representations of phenomena has been shown to support students to retain and integrate complex knowledge (Duschl et al., 2007). Yet, many existing elementary science curriculum materials may be grounded in theoretical perspectives and assumptions about student learning and development that do not reflect empirical findings from contemporary education research or fundamental tenets of science education reform (Duschl et al., 2007; NRC, 2012). Student sense making

through engagement in explanation construction must be foregrounded in the science curriculum materials made available to elementary teachers at scale, particularly if they are not being significantly adapted by teachers who use them.

Second, elementary teachers at both the in-service and preservice stages along the teacher professional continuum require substantial and meaningful opportunities to develop robust understanding of explanation construction AND how to effectively support students to formulate evidence-based explanations. One crucial means to provide this support is through the development of educative curriculum materials (Davis & Krajcik, 2005) that support teachers to engage students in explanation construction by providing embedded features that, when accessed by teachers using the curriculum materials, can promote teachers' learning and practice. While the majority of the curriculum materials examined in this study did include content support for the teachers, they included limited educative elements to support teachers' understanding of how to support students' scientific sense making or possible student misconceptions. This supports findings from reviews of middle and secondary science curriculum materials that include few meaningful educative elements for teachers (Beyer et al., 2009; Kesidou & Rosemann, 2002). Educative supports designed around the cause, effect, and mechanism perspective on explanation construction (Braaten & Windschitl, 2011; Chinn & Malhotra, 2002; Windschitl et al., 2008) and contemporary frameworks for scientific practices (NRC, 2000, 2012) could highlight those lesson components that engage students in explanation construction and differentiate between what happens (cause and effect) and why or how (mechanism). They would explicitly identify transition points in unit lessons in which students move from data analysis, or representing cause and effect, to formulating mechanism-based claims for their observations. By tying lesson elements to components of underlying conceptual frameworks, teachers would be provided rationales for particular instructional approaches, a fundamental design heuristic for educative curriculum materials (Davis & Krajcik, 2005).

A sustained, long-term program of support is necessary for teachers to productively foster elementary science learning environments centered on students' formulation of evidence-based explanations. A change in teaching practices requires teachers' to experience a conceptual change in their beliefs about student learning (Beyer & Davis, 2008; Metz, 2009). In order to change beliefs, experienced teachers must engage in a reform effort and then see a positive response that demonstrates their students are capable of successfully engaging in the reform. However, as

Metz (2009) shows, even with extensive professional development and well-developed curriculum, an experienced elementary teacher who enters into the reform with limited disciplinary knowledge will have difficulty conceptualizing the necessity for implementing the reforms and, as such, will not include it in her enactment. We suggest that a crucial component of effective science teacher education and professional development experiences for elementary teachers that foreground students' explanation construction is a firm grounding in effectively-designed elementary science curriculum materials beginning in a teacher's preservice education (Avraamidou & Zembal-Saul, 2005; Forbes & Davis, 2010). But, as findings from this study suggest, many curriculum-based elementary-level science investigations used in daily practice may not afford robust sense making opportunities. If students are to be afforded opportunities for sense making, then teachers must modify their instruction to better support students' explanation construction. Prior research with preservice teachers shows that they are able to make effective adaptations to science curriculum materials to better promote students' sense making about natural phenomena (Forbes & Davis, 2010; Zangori & Forbes, 2013). Within the context of explanation construction as a core feature of inquiry (NRC, 2000) or scientific practice (NRC, 2012), the cause, effect, mechanism framework for explanation construction (Braaten & Windschitl, 2011; Windschitl et al., 2008) could serve as a powerful heuristic for teachers' evaluation and adaptation of existing elementary science curriculum materials. Over the long term, and through collaboration with peers using the same curriculum materials for science, teachers can learn to employ their pedagogical reasoning effectively to make instructional decisions about how to best foster explanation construction in their own classrooms.

Conclusion

This research adds to the limited body of research on in-service elementary teachers' classroom practices and use of science curriculum materials. This work supports and extends the notion that, as with beginning elementary teachers, experienced elementary teachers learning about explanation construction requires support. First, future research should explore elementary teachers' knowledge about formulating explanations and how they interpret and interact with their curricular materials over time to afford elementary students opportunities to formulate explanations in the classroom. This work is particularly needed to understand

the impact of particular interventions—whether educative curriculum materials, professional development, or both—on changes in teachers’ knowledge and practices. Second, we did not measure student learning as part of this study. Future work should explore relationships between patterns in teachers’ curriculum adaptation and student outcomes. Such work will help science teacher educators and curriculum developers better understand how to support teachers to foster explanation construction in elementary science learning environments.

Contract grant sponsors: Roy J. Carver Charitable Trust and University of Iowa College of Education.

References

- Avraamidou, L., & Zembal-Saul, C. (2005). Giving priority to evidence in science teaching: A first-year elementary teacher’s specialized practices and knowledge. *Journal of Research in Science Teaching*, 42(9), 965–986. doi: 10.1002/tea.2081
- Berland, L., & Reiser, B. J. (2008). Making sense of argumentation and explanation. *Science Education*, 93(1), 25–55. doi: 10.1002/sce.20286
- Berland, L. K., & McNeill, K. L. (2012). For whom is argument and explanation a necessary distinction? A response to Osborne and Patterson. *Science Education*, 96(5), 808–813.
- Beyer, C. J., & Davis, E. A. (2008). Fostering second graders’ scientific explanations: A beginning elementary teacher’s knowledge, beliefs, and practice. *Journal of the Learning Sciences*, 17(3), 381–414. doi: 10.1080/10508400802222917
- Beyer, C. J., Delgado, C., Davis, E. A., & Krajcik, J. (2009). Investigating teacher learning supports in high school biology curricular materials. *Journal of Research in Science Teaching*, 46(9), 977–998. doi: 10.1002/tea.20293
- Biggers, M., Forbes, C. T., & Zangori, L. (in press). Elementary teachers’ curriculum design and pedagogical reasoning for supporting students’ comparison and evaluation of evidence-based explanations. *The Elementary School Journal*.
- Braaten, M., & Windschitl, M. (2011). Working toward a stronger conceptualization of scientific explanation for science education. *Science Education*, 95(4), 639–669.
- Cavagnetto, A., Hand, B. M., & Norton-Meier, L. (2010). The nature of elementary student science discourse in the context of the science writing heuristic approach. *International Journal of Science Education*, 32(4), 427–449. doi: 10.1080/09500690802627277

- Chinn, C. A., & Malhotra, B. A. (2002). Epistemologically authentic inquiry in schools: A theoretical framework for evaluating inquiry tasks. *Science Education*, 86(2), 175–218. doi: 10.1002/sce.10001
- Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research*. Thousand Oaks, CA: SAGE Publications.
- Davis, E. A., & Krajcik, J. (2005). Designing educative curriculum materials to promote teacher learning. *Educational Researcher*, 34(3), 3–14. doi: 10.3102/0013189X034003003
- Duschl, R. (2008). Science education in three-part harmony: Balancing conceptual, epistemic, and social learning goals. *Review of Research in Education*, 32(1), 268–291.
- Duschl, R. A., Schweingruber, H. A., & Schouse, A. W. (Eds.). (2007). *Taking science to school: Learning and teaching science in grades K-8*. Washington, DC: National Academy Press.
- Forbes, C. T., Biggers, M., & Zangori, L. (2013). Investigating essential characteristics of scientific practices in elementary science learning environments: The Practices of Science Observation Protocol (PSOP). *School Science and Mathematics*, 4(113), 180–190.
- Forbes, C. T., & Davis, E. A. (2010). Curriculum design for inquiry: Preservice elementary teachers' mobilization and adaptation of science curriculum materials. *Journal of Research in Science Teaching*, 47(7), 820–839. doi: 10.1002/tea.20379
- Hagood, S., Magnusson, S. J., & Palinscar, A. S. (2004). Teacher, text, and experience: A case of young children's scientific inquiry. *The Journal of the Learning Sciences*, 13(4), 455–505.
- Hardy, I., Jonen, A., Möller, K., & Stern, E. (2006). Effects of instructional support within constructivist learning environments for elementary school students' understanding of "floating and sinking". *Journal of Educational Psychology*, 98(2), 307–325.
- Kesidou, S., & Rosemann, J. E. (2002). Do middle school science programs measure up? Findings from project 2061's curriculum review. *Journal of Research in Science Teaching*, 39(6), 522–549. doi: 10.1002/tea.10035
- King, K., Shumow, L., & Lietz, S. (2001). Science education in an urban elementary school: Case studies of teacher beliefs and classroom practices. *Science Education*, 85(2), 89–110. doi: 10.1002/1098-237X(200103)85:2<89:AID-SCE10>3.0.CO;2-H
- Littell, R. C., Milliken, G. A., Stroup, W. W., Wolfinger, R. D., & Schabenberger, O. (2006). *SAS for mixed models* (2nd ed.). Cary, NC: SAS Institute, Inc.
- Magnusson, S., & Palinscar, A. M. (2005). Teaching to promote the development of scientific knowledge and reasoning about light at the elementary school level. In M. S. Donovan, & J. D. Bransford (Eds.), *How students learn: Science in the classroom* (pp. 421–474). Washington, DC: National Academies Press.
- McNeill, K. L. (2011). Elementary students views of explanation, argumentation, and evidence, and their abilities to construct arguments over the school year. *Journal of Research in Science Teaching*, 48(7), 793–823. doi: 10.1002/tea.20201

- McNeill, K. L., & Krajcik, J. (2008). Scientific explanations: Characterizing and evaluating the effects of teachers' instructional practices on student learning. *Journal of Research in Science Teaching*, 45(1), 53-78. doi: 10.1002/tea.20201
- Metz, K. (1995). Reassessment of developmental constraints on children's science instruction. *Review of Educational Research*, 65(93), 93-127. doi: 10.3102/00346543065002093
- Metz, K. (2009). Elementary school teachers as "targets and agents of change": Teachers' learning in interaction with reform science curriculum. *Science Education*, 93(5), 915-954. doi: 10.1002/sce.20309
- Metz, K. (2011). Disentangling robust developmental constraints from the instructionally mutable: Young children's epistemic reasoning about a study of their own design. *Journal of the Learning Sciences*, 20 (1), 50-110. doi: 10.1080/10508406.2011.529325
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis*. Newbury Park, CA: Sage.
- Minogue, J., Madden, L., Bedward, J., Wiebe, E., & Carter, M. (2010). The cross-case analysis of elementary students' engagement in the strands of science proficiency. *Journal of Science Teacher Education*, 21(5), 559-587. doi: 10.1007/s10972-010-9195-y
- National Research Council. (2000). *Inquiry and the national science education standards: A guide for teaching and learning*. Washington, DC: The National Academy Press.
- National Research Council. (2012). *A framework for K-12 science education: Practices, cross-cutting concepts, and core ideas*. Washington, DC: National Academies Press.
- Osborne, J., & Patterson, A. (2011). Scientific argument and explanation: A necessary distinction? *Science Education*, 95(4), 627-638. doi: 10.1002/sce.20438
- Patton, M. Q. (2001). *Qualitative research and evaluation methods*. Thousand Oaks, CA: Sage.
- Remillard, J. T. (2005). Examining key concepts in research on teachers' use of mathematics curricula. *Review of Educational Research*, 75(2), 211-246.
- Ruiz-Primo, M. A., Li, M., Tsai, S., & Schneider, J. (2010). Testing one premise of scientific inquiry in science classrooms: Examining students scientific explanations and student learning. *Journal of Research in Science Teaching*, 47(5), 583-608.
- Salmon, W. (1998). *Causality and explanation*. New York, NY: Oxford University Press.
- Samarapungavan, A., Mantzicopoulos, P., & Patrick, H. (2008). Learning science through inquiry in kindergarten? *Science Education*, 92(5), 868-908. doi: 10.1002/sce.20275
- Songer, N. B., & Wenk Gotwals, A. (2012). Guiding explanation construction by children at the entry points of learning progressions. *Journal of Research in Science Teaching*, 49(2), 141-165. doi: 10.1002/tea.20454

- Trout, J. D. (2002). Scientific explanation and the sense of understanding. *Philosophy of Science*, 69(2), 212–233. doi: 10.1086/341050
- Winschitl, M., Thompson, J., & Braaten, M. (2008). Beyond the scientific method: Model-based inquiry as a new paradigm of preference for school science investigations. *Science Education*, 92(5), 941–967.
- Yin, R. K. (2009). *Case study research: Design and methods*. Thousand Oaks, CA: Sage.
- Zangori, L., & Forbes, C. T. (2013). Preservice elementary teachers and explanation construction: Knowledge-for-practice and knowledge-in-practice. *Science Education*, 97(2), 310–330.
- Zangori, L., Forbes, C. T., & Biggers, M. (2012). This is inquiry ... right? Strategies for effectively adapting elementary science lessons. *Science and Children* 50(1), 48–53.