Elementary Teachers' Use of Formative Assessment to Support Students' Learning About Interactions Between the Hydrosphere and Geosphere

Cory T. Forbes,[1,2,a] Jaime L. Sabel,[2] and Mandy Biggers[3]

ABSTRACT

Students' thinking should serve as the foundation of effective science curriculum and instruction. To promote science learning, particularly in the geosciences, teachers must attend to students' existing ideas about natural phenomena through the use of ''high-leverage'' instructional practices such as formative assessment. Elementary teachers need support to learn to implement formative assessment practices effectively. However, few studies have explored relationships between elementary teachers' content knowledge and formative assessment practices, analysis of students' thinking, and instructional decision making. To begin to address this gap in the literature, we conducted a convergent parallel mixed methods study to examine how elementary teachers employ formative assessment practices to scaffold elementary students' learning about interactions between water and the geosphere, which is a core, unifying concept in the Earth sciences. This research is embedded within a multiyear professional development program designed to support elementary teachers (grades 3–5) to learn to employ formative assessment in their classrooms. Study findings show teachers' own knowledge of geoscience disciplinary content is unrelated to their formative assessment practices. They also highlight the importance of the elementary science curriculum materials teachers use in framing disciplinary concepts in ways that influence how teachers evaluate student artifacts and engage in follow-up instruction based. © *2015 National Association of Geoscience Teachers*. [DOI: 10.5408/14-063.1]

Key words: geoscience, elementary science, formative assessment, elementary teachers

INTRODUCTION

Elementary science is crucial to help lay the foundation for students' lifelong science learning (National Research Council [NRC], 2007). Toward that end, science standards emphasize the need for students to develop robust understanding of Earth's systems, including the geosphere and hydrosphere (Next Generation Science Standards [NGSS] Lead States, 2013). However, early learners possess a range of ideas about water and how it interacts with the Earth, many of which are inconsistent with scientific knowledge (Bar, 1989; Cheek, 2010; Forbes et al., 2015). Students therefore need support through instruction to develop scientifically accurate explanations for these phenomena.

Student learning can be fostered through the use of recognized ''high-leverage'' instructional strategies (Ball and Forzani, 2009). Formative assessment is an established and proven high-leverage instructional practice in which teachers use evidence of students' thinking to shape instruction and, thus, can better support students' scientific sense making. Use of formative assessment practices within classrooms has been shown to lead to significant student learning gains in science (Ruiz-Primo and Furtak, 2006). However, the use of formative assessment for science is not widespread in elementary classrooms (Coffey et al., 2011; Morrison, 2013). As has been shown with other instructional interventions at the elementary level (e.g., Hill et al., 2005), teachers must possess strong knowledge of the disciplinary content if students are to learn to implement formative assessment practices effectively (Brookhart, 2011). However, elementary teachers are often ill prepared to teach disciplinary content, particularly in the geosciences (Marcum-Dietrich et al., 2011) and little is known about how they leverage their disciplinary knowledge and navigate curricular and instructional affordances and constraints to engage in formative assessment for science. While learning to engage in formative assessment itself can serve as ''effective and transformative professional development for teachers'' (Ash and Levitt, 2003, 44), elementary teachers need sustained support in learning to integrate formative assessment practices into their instructional repertoire.

To address these needs, we provide support for third- through fifth-grade teachers to learn to implement effective formative assessment practices through a multiyear professional development program designed around key features of effective professional development for science teachers (e.g., Penuel et al., 2009; Heller et al., 2012). The objective of this program is to support teachers' learning to integrate and implement formative assessment practices into their science instruction. To support the ongoing development of the program, as well as to contribute to the knowledge base in science education, we are also engaged in related research around teachers' implementation of formative assessment practices. Both the program and associated research are grounded in the hypothesis that teachers with stronger science content knowledge will be better positioned to engage more effectively in formative assessment practices. In

Received 18 November 2014; accepted 7 April 2015; published online 14 September 2015.
[1]School of Natural Resources, University of Nebraska–Lincoln, 523 Hardin Hall, 3310 Holdrege Street, Lincoln, Nebraska 68583-0995, USA
[2]Department of Teaching, Learning, and Teacher Education, College of Education and Human Sciences, University of Nebraska–Lincoln, 233 Mabel Lee Hall, Lincoln, Nebraska 68588, USA
[3]College of Education, The Pennsylvania State University, 226 Chambers Building, University Park, Pennsylvania 16802-0234, USA
[a]Author to whom correspondence should be addressed. Electronic mail: cory.forbes@unl.edu. Tel.: 402-472-7844. Fax: 402-472-2946.

this paper, we report on findings from the first year of the program, which focused on supporting students' learning in the geosciences, to answer the following research questions:

1. To what extent does teachers' disciplinary knowledge of Earth Science influence their ability to engage in formative assessment practices?
2. How do teachers identify trends in students' thinking and implement follow-up instruction about earth science concepts?

Significance

While prior studies have explored preservice and inservice elementary teachers' use of formative assessment for science (e.g., Morrison and Lederman, 2003; Aschbacher and Alonzo, 2006; Otero and Nathan, 2008; Buck and Trauth-Nare, 2009; Coffey et al., 2011; Hammer et al., 2012; Morrison, 2013; Sabel et al., 2015), additional research is needed to address key gaps that are still present in existing research. Specifically, research is needed to illustrate (1) relationships between teachers' disciplinary knowledge and formative assessment practices and (2) the resources teachers mobilize to inform their diagnosis of students' thinking and instructional decision making. This study addresses these identified gaps and contributes to a body of work focused on teaching and learning in the geosciences (Bar, 1989; Cheek, 2010; Forbes et al., 2015). The study has important implications for K–6 Earth Science instruction, research on teachers' reasoning and instructional practices, and the design of professional development and science curriculum materials.

BACKGROUND AND THEORETICAL FRAMEWORK

Theoretical Framework—Formative Assessment and Science Learning

To support student learning about water- and Earth Science-related phenomena, elementary science learning environments should be designed to embody core tenets of contemporary learning theory. Formative assessment is an instructional practice through which teachers can craft learning environments in ways that are responsive to students' ideas. The classroom environment influences both the students' scientific reasoning and the teachers' pedagogical reasoning and, in turn, is influenced by these factors. In the classroom context, teachers elicit, interpret, and diagnose their students' ideas and then plan revised experiences in which the students will engage to help develop their pre-existing ideas into new or revised ideas. As such, formative assessment aligns directly with current views on science learning and can directly impact students' developing understanding of natural phenomena. A conceptual framework for this relationship is shown in Fig. 1.

Current reform-based views on learning revolve around principles that underlie constructivism (e.g., Donovan and Bransford, 2005), which place the learner in the active role of building new knowledge through engagement in meaningful activity rather than as a passive recipient of information. A core assumption of constructivist learning theories is that students already possess ideas about the natural world that may be sensible but not scientifically accurate or complete. In effectively designed science learning environments, students' ideas serve as the currency of the classroom because children need to recognize, build on, and refine their ideas (NRC, 2007). Through participation in domain-specific practices, students learn to generate new knowledge through processes and norms that define particular sense-making communities. The intellectual and practical work associated with interrogating and refining science ideas over time is pursued through engagement in scientific practices such as questioning, investigation, explanation construction, and argument (NGSS Lead States, 2013).

Such sense making involves both individual and social processes and activities. Teachers are also part of the classroom communities to which students belong. Effective teachers provide meaningful opportunities for students to articulate their thinking and pay close attention to trends and patterns in students' ideas. This information informs teachers' pedagogical reasoning, through which they make decisions about appropriate plans for ongoing instruction. Additionally, teachers, like students, are afforded opportunities to develop knowledge through engagement in effective, reform-based instruction that is responsive to students' ideas. Teachers who engage in science instruction that is responsive to students' ideas can learn from those ideas as they respond to them within classroom communities comprised of both teachers and students (Ash and Levitt, 2003). As such, elementary science learning environments grounded in students' thinking provide rich contexts for both teaching and student learning, including within Earth Science contexts.

Formative Assessment

Formative assessment represents a set of instructional practices and specific strategies that embody the theoretical underpinnings of student-responsive instruction. Formative assessment allows teachers to account for individual student progress, craft responsive instruction, and cultivate student-centered science learning environments (Bell and Cowie, 2001; Coffey et al., 2011). Teachers can assess students' understanding and adjust instruction to engage students more effectively in practices that promote sense making and scientific learning. Past research suggests that teachers' use of formative assessment to inform their instruction has led to significant gains in student learning (Ruiz-Primo and Furtak, 2006).

Formative assessment practices can vary along a continuum from "on-the-fly" to "planned-for" interaction to "curriculum-embedded" assessment (Shavelson et al., 2008). On-the-fly formative assessment practices do not involve a planned activity, but rather arise out of the learning activity when a teacher recognizes and responds to student thinking (Cowie and Bell, 1999). This may occur when a teacher overhears students discussing an idea and responds to it in the moment. Planned-for formative assessment occurs when a teacher deliberately plans to stop and check for student understanding and to respond to it as needed. Teachers may utilize prompts or student artifacts to invite students to evaluate their ideas. Finally, embedded formative assessment includes predefined tasks built into the curriculum that can help teachers locate their students' ideas and provide accurate feedback (Ayala et al., 2008). These assessments are typically placed at specific junctures in curricular sequences and are used to determine if students have learned an important subgoal of the unit. Through

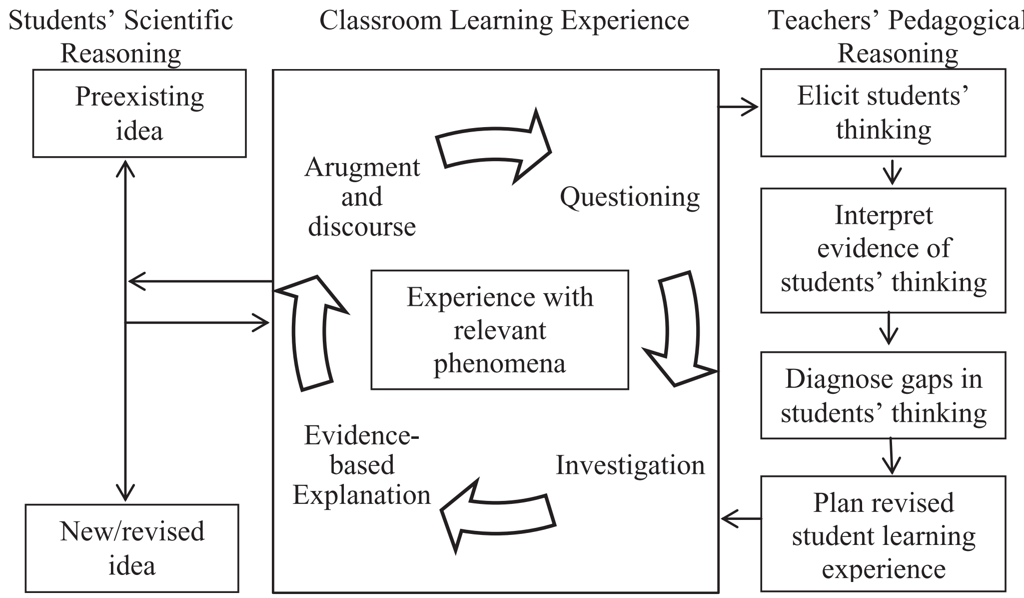

FIGURE 1: Relationship between formative assessment and students' science learning (Sabel et al., 2015).

these approaches, formative assessment provides teachers with information they can use to engage in science instruction that is responsive to students' ideas.

Research on Formative Assessment for Elementary Science

In order to use instruction to support their students' sense making, teachers should "know how students think, have strategies for eliciting their thinking as it develops, and use their knowledge flexibly in order to interpret and respond strategically to student thinking" (NRC, 2007, 312). Teachers' abilities to respond appropriately to students' ideas and provide support for student engagement in scientific practices are crucial to help students move toward established scientific knowledge. However, both preservice and inservice teachers often don't understand what formative assessment is, know how to employ it in the classroom, or possess sufficient knowledge of content to do so effectively (Otero and Nathan, 2008; Coffey et al., 2011; Hammer et al., 2012; Sabel et al., 2015). These factors are further complicated by the use of instructional materials for elementary science, which may or may not reflect contemporary perspectives on student learning and facilitate student responsive instructional approaches (Buck and Trauth-Nare, 2009; Heritage et al., 2009). Elementary teachers tend to rely heavily on their curricular resources for science (Biggers et al., 2013; Forbes et al., 2013; Zangori et al., 2013), which can present both opportunities and challenges for teachers engaging in instruction that is responsive to students' ideas. All of these factors combined point to the importance of sustained support to help elementary teachers learn to integrate formative assessment into their instructional repertoire, particularly for science.

Prior research has highlighted the importance of disciplinary knowledge for elementary teachers, showing that teachers who possess robust content knowledge are better able to implement more effective teaching strategies and to support students' learning (Hill et al., 2005; Falk, 2012). Findings from a small number of studies similarly suggest that elementary teachers' content knowledge can influence their implementation of formative assessment (Aschbacher and Alonzo, 2006; Heritage et al., 2009; Coffey et al., 2011), reinforcing a core assumption of teacher preparation standards (Brookhart, 2011). Content knowledge is important not only for effectively eliciting and diagnosing students' thinking (Morrison and Lederman, 2003; Gottheiner and Siegel, 2012), but also for implementing discipline-specific instructional strategies that scaffold students' learning. Taken together, these findings suggest that increasing teachers' disciplinary knowledge is likely to enhance their instruction, including the use of formative assessment, which could translate to greater student achievement gains. However, few studies have yet to explore relationships between elementary teachers' science content knowledge and formative assessment practices. This study is intended to fill this gap in the research with a focus on geoscience content knowledge. Consistent with past research and conventional wisdom in the field, we hypothesize that teachers with stronger knowledge of disciplinary content, in this case for the Earth sciences, are able to engage more productively in formative assessment practices.

STUDY DESIGN AND METHODS

The Professional Development Program

This study was conducted during the first year of a 3-year sustained professional development program to support third- through fifth-grade teachers to implement the Reflective Assessment (RA; Kennedy et al., 2009) formative assessment strategy in science. RA is an example of planned-for formative assessment comprised of a four-step process in which teachers *anticipate* students' ideas, *teach* a lesson, *review* student artifacts, *reflect* on patterns in student thinking, and *adjust* instruction based on their diagnosis of students' thinking. The professional development program was designed around core features of effective science professional development (Penuel et al., 2009; Marcum-Dietrich et al., 2011; Heller et al., 2012), such as being localized and curriculum-specific, long-term and sustained, and integrating Earth Science content and pedagogy. Teachers participated in a 7-day summer institute before the academic year in which study data was collected. During the school year, they also took part in bimonthly meetings of teacher collaborative learning teams and two half-day workshops, all focused on the Earth Science units the teachers already used. Program activities focused on supporting teachers' learning of Earth Science content, implementation of formative assessment to support students' learning about interactions between the hydrosphere and geosphere, examination of student work and artifacts of teacher practice, developing cases from their own classrooms, performing Earth Science investigations as learners, using evidence to anticipate likely student misconceptions about Earth Science concepts, and implementing preidentified instructional strategies in ways that were responsive to students' thinking.

Context and Participants

The project involved 26 third- through fifth-grade teachers recruited from 13 schools in three school districts in a Midwestern state. Twenty-one of these teachers were female and five were male, and their teaching experience ranged from three years to more than 20 years. These teachers were selected from a larger pool of teachers using kit-based instruction materials from the Full Option Science System (FOSS) and other commercial publishers for science, all provided through a regional science curriculum center. All teachers taught at least one Earth Science unit that foregrounded the hydrosphere and geosphere, emphasizing topics such as properties of water, surface and subsurface flow, and erosion and deposition. These Earth Science units each lasted approximately 8 weeks and were taught at various times throughout the school year depending on each teacher's instructional assignment and curricular schedule.

Out of these 26 teachers, six were purposefully selected (Table I; Merriam, 2009) to participate in embedded case studies. This selection was based on the particular Earth Science units they were teaching, their previous collaboration on unit planning during the summer professional development workshop, and their representative scores on the assessment of Earth Science content knowledge. Three third-grade teachers, Anna, Sheri, and Kim, taught the FOSS Water Module (Full Option Science System [FOSS], 2005a), while Alicia, Cathryn, and Matt, all fifth-grade teachers, taught the FOSS Landforms Module (FOSS, 2005b).

Data Collection

We used a convergent parallel mixed methods research design (Creswell and Plano Clark, 2011) to investigate how teachers examine student work and select follow-up instructional steps to address the trends in student understanding they identify. This approach allowed us to collect multiple forms of data in parallel, utilize both quantitative and qualitative data analysis methods, and then merge findings to reach greater depth of understanding of how teachers engaged in formative assessment (Creswell and Plano Clark, 2011). To address the first research question, we used quantitative methods to examine relationships between the science content knowledge and formative assessment use of all 26 teachers engaged in the project. We also used qualitative methods to conduct a multiple-case study with cross-case analysis (Miles and Huberman, 1994; Yin, 2009) that focused on the instructional decisions and practices that six teachers used when implementing formative assessment in their classrooms to address research question 2.

The project team developed instructional logs for all 26 teachers to use to document their use of RA, the student work they examined, the trends in students' ideas they observed, and details on the rationale for the follow-up instruction they used. Instructional logs have previously been used in similar classroom-based research contexts to characterize instruction through self-report (Camburn and Barnes, 2004; Rowan and Correnti, 2009). Log questions were designed to mirror components of the RA strategy (Table II). Logs included both open-ended and forced-response items that allowed teachers to document their use of RA, as well as articulate their own pedagogical reasoning. Teachers accessed the logs in an online format from their classroom computers and submitted them electronically to the research team. On average, teachers required 16 min to complete each log. Along with each log, teachers also submitted hard copies of student work samples and instructional artifacts from the lesson on which each log was based. The logs, sample student work, and other documents were compiled into a single file and stored electronically. Teachers were asked to submit 12 logs during the school year (approximately four logs for each science unit they taught). Additionally, some of the teachers taught more than one Earth Science unit each year. So that teachers with more than one unit were not overrepresented, we limited the data to only include logs from a single Earth Science unit per teacher. The logs submitted by all 26 teachers for their Earth Science units ($n = 73$) were used as data for this study.

All 26 teachers in the project also completed an assessment of their knowledge of science content that aligned with the Earth Science units they taught. Items were selected from the Misconceptions-Oriented Standards-Based Assessment Resources for Teachers (MOSART) for grades five through eight, which was designed to be used with teachers or students (Sadler et al., 2010). The MOSART is comprised of multiple-choice items that align with both science standards and research on students' misconceptions. MOSART assessments have been shown to be reliable and valid standards-based measures of students' and teachers' knowledge (Sadler et al., 2010, 2013; Wendt and Rockinson-Szapkiw, 2014). We chose to use the items to assess the teachers' content knowledge at a slightly higher level than what they teach. The assessment consisted of 40 items, including 12 Earth Science items that were selected to specifically align with geoscience concepts related to content in the teachers' Earth Science units around interactions between the hydrosphere and geosphere. The measure of teachers' Earth Science content knowledge used here is based upon the subset of 12 Earth Science items from the assessment instrument. The raw mean interitem correlation for these 12 Earth Science items is 0.14. This value falls on the edge of the optimal range for internal consistency recommended by Clark and Watson (1995) for measure of broad, top-level constructs (such as Earth Science content knowledge) using a limited number of items with a relatively small sample.

In addition to the instructional logs and assessment data collected from all teachers, additional data were collected from a subset of six third- and fifth-grade teachers. First, we interviewed each of these six teachers at the beginning and end of the unit using a semistructured interview protocol (Merriam, 2009). We also observed and video-recorded classroom enactments of unit lessons in which teachers engaged in RA, including follow-up instruction. Each observation was video-recorded and followed by with a lesson-specific, semistructured interview (Table II). This resulted in at least four observations and four lesson-specific interviews for each teacher. These interviews, video-recorded lesson enactments, and logs were used to build case studies (Yin, 2009) that examined how the teachers used student work to identify students' thinking about water and earth materials and how the teachers used those student ideas to make decisions regarding follow-up instruction. All participants were assigned an anonymous code. The interviews were transcribed verbatim. Teachers' logs and associated artifacts were scanned and saved digitally. All data was imported into ATLAS.ti. The video-recorded classroom observations were analyzed for supporting and disconfirming evidence of the emergent themes identified in interviews and logs.

Scoring and Quantitative Analysis

All Earth Science lesson logs were scored by two scorers using a rubric we developed to examine teachers' engagement in RA. The rubric provided a tool to evaluate how effectively teachers were engaging in individual components of RA. These included (1) defining the key concept and what they were looking for in student responses (*anticipate*), (2) explaining trends in student understanding (*review*/*reflect*), and (3) providing rationale for their choice of follow-up instruction (*adjust*). Each log was scored on a five-point scale that ranged from 0 to 4 (5 levels). Scoring levels were defined in terms of variations in effective practices associated with RA. A summary of typical high-scoring and low-scoring responses on the lesson logs is provided in Table III. An example of the rubric for the RA component of *adjust* is included in the Supplemental Material A, available online at http://dx.doi.org/10.5408/14-063.1s1.The rubric was vetted by RA developers and went through multiple rounds of revision until two scorers obtained consistent scores. We calculated the intraclass correlation in SPSS as a measure of interscorer reliability by comparing the scorers' average score on each lesson log using a two-way mixed model ANOVA for consistency. The intraclass correlation coefficient for the two scorers across all Earth Science lesson logs was 0.808.

We examined the differences between the log scores from teachers' Earth Science units and the assessment scores using a mixed-model analysis in SAS. Each teacher submitted more than one lesson log in a sequential manner (nested per teacher), but had only a single content exam score; therefore, our analysis required us to use a multilevel mixed model to account for the repeated measures and nesting. We performed a single-factor repeated-measures mixed model ANOVA (Littell et al., 2006), in which the dependent variable was the average log score for each of the Earth Science unit logs for each teacher and the independent variable was the content exam score for each teacher. The single-factor, repeated-measures, mixed model ANOVA formula is $Y_{ij} = \gamma_{00} + \gamma_{01} \times z_j + U_{0j} + R_{ij}$, where Y_{ij} is the score for each teacher (j) on each log (i); γ_{00} is the grand mean; γ_{01} is the effect for the total score; z_j is the total score for each teacher (j); U_{0j} is the random effect for each teacher (j); and R_{ij} is the random error for each teacher (j) on each log (i). Finally, to examine the correlation between teachers' specific content knowledge pertaining to the lessons studied here, we also performed one-way ANOVA (Littell et al, 2006) for each of the individual questions on each lesson log. Here, the dependent variable was the content exam score and the independent variable for each test was an individual question on the teachers' lesson logs.

Qualitative Data Coding and Analysis

We used multiple-case study methodology with cross-case analysis to examine artifacts from the six case-study teachers (Miles and Huberman, 1994; Yin, 2009). Analysis of multiple sources of data allows our research questions to be viewed from various angles (logs, videos, and interviews) and provides triangulation of our findings (Merriam, 2009). All 34 interviews and 20 Earth Science logs submitted by the case study teachers (Table I) were coded using a priori codes established from the RA cycle: *anticipate*, *teach*, *review*, *reflect*, and *adjust* (Kennedy et al., 2009; Supplemental Material B, available online at http://dx.doi.org/10.5408/14-063.1s2). We established intercoder reliability between two coders by coding a 10% sample of the data. Interrater reliability was approximately 71% before discussion and 100% following discussion. We then queried the *review*, *reflect*, and *adjust* codes. In our second round of coding, we analyzed the data in these three first-round codes through open-coding (Merriam, 2009) to identify emergent codes (Supplemental Material B) and common themes across the data. This emergent coding resulted in the themes we present in the findings. The video-recorded classroom observations were examined using the themes that emerged from data coding to provide confirming and/or disconfirming evidence. These themes were compared using a pattern-matching strategy (Yin, 2009) to find common patterns across all six teachers.

RESULTS

Relationship Between Teachers' Content Knowledge and Formative Assessment Practices

In the first research question, we asked, "To what extent does teachers' disciplinary knowledge of Earth Science influence their ability to engage in formative assessment practices?" We hypothesized that teachers with more robust knowledge of Earth Science concepts would be more effective at evaluating students' ideas and engaging in subsequent, student-responsive instruction. Descriptive statistics for the content exam and lesson log scores are presented in Table IV. Results of statistical analysis yielded no observable relationships between teachers' content knowledge and the quality of their formative assessment practices. As shown in Fig. 2, there is no relationship between the teachers' Earth Science content exam score and their average lesson log scores.

We observed no statistically significant relationship between teachers' Earth Science content knowledge, as measured by the Earth Science item score, and their overall lesson log score, $F(3, 70) = 0.08$, $p = 0.78$. We also compared the teachers' Earth Science content exam scores against the

TABLE I. Summary of case study teachers and classrooms.

	Anna	Sheri	Kim	Alicia	Cathryn	Matt
Grade level	3rd	3rd	3rd	5th	5th	5th
FOSS unit	Water	Water	Water	Landforms	Landforms	Landforms
Class size	22	21	11	27	30	31
Years of teaching experience	21	22	23	23	15	3
School distribution	School 1	School 1	School 2	School 3	School 3	School 4
School % eligible free or reduced lunch	24.5%	24.5%	4.3%	26.6%	26.6%	54.3%
Earth Science logs submitted	3	3	5	4	3	3
Interviews	6	6	6	6	5	5
Content exam score (out of 12)	8	7	10	8	7	10

scores for each component of the RA cycle (*anticipate*, *review/reflect*, *adjust*) on the lesson logs (see Table IV for descriptive statistics). Again, our results indicate no statistically significant relationships between teachers' Earth Science content knowledge and any of the constituent components of RA measured in the lesson log analysis (data not shown).

Teachers' Evaluation of Students' Ideas and Instructional Decision Making

Although we had hypothesized that teachers with stronger content knowledge would engage more productively in formative assessment practices, results of quantitative analysis of logs from this group of teachers did not support this assumption. Therefore, we turned to qualitative analysis to further explore how the teachers were engaging in formative assessment practices if not by leveraging their knowledge of Earth Science content. We utilize results from qualitative analyses to answer research question 2, in which we asked, ''How do teachers identify trends in students' thinking and implement follow-up instruction about Earth Science concepts?'' Rather than draw upon their own knowledge of Earth Science concepts, teachers relied heavily on the curriculum materials to define the disciplinary nature of their students' ideas. They used student artifacts, typically provided by the curriculum, to evaluate students' ideas and looked for specific curriculum-defined markers, such as vocabulary terms. In effect, they offloaded much of the conceptual framing of target Earth Science concepts for students to the curriculum modules the used. As a result, when they did engage in follow-up instruction based upon a set of available discipline-nonspecific instructional strategies that were not part of the curriculum modules, the teachers tended to use the same strategies repeatedly rather than employ follow-up instructional strategies aligned with gaps in student understanding they identified. The representative quotes and snapshots from classroom observations we present serve as exemplars for broader trends in the data.

Interpreting and Evaluating Student Work

In their analysis of students' thinking, teachers typically emphasized students' use of particular terminology and other representations that aligned with the ways individual unit lessons operationalized target concepts for students. They tended to define conceptual goals for students in terms of the curriculum, focusing on words such as *density* or *surface tension* (for the Water unit) or *erosion* and *deposition* (for the Landforms unit). For example, in one of her lessons, Alicia described her key concept as, ''exactly what the class

TABLE II. Lesson log and interview questions aligned with the reflective assessment framework.

Steps of Reflective Assessment	Lesson Log Questions	Examples of Interview Questions
Anticipate/teach	• What is the key concept you are assessing in the student work?	• What was the key concept you chose to look at for this log? • Where did that key concept come from? • What makes it a good key concept to look for?
Review	• What student work are you examining? • What are you looking for in students' responses as evidence of understanding of the key concept? • How many of your students got it/still need help?	• What types of student work did you look at? • Why did you choose that particular student work sample?
Reflect	• What did you notice in the student work you reviewed? • Did you observe any misconceptions? • What trends in students' understanding of the key concept did you see?	• What did you notice as you looked through the students' work? • What did you see as evidence of student understanding? • What were examples of student misconceptions?
Adjust	• Why did you select this follow-up instructional strategy? • How did you think it will enhance students' understanding of the key concept?	• What follow-up instructional strategy did you use? How did that go? • Was the follow-up instruction effective? • Would you do anything differently the next time you taught the content?

TABLE 3. Comparison of high scorers and low scorers on the lesson logs.

Lesson Log Section	High Scorers	Low Scorers
Anticipate/teach	• Identified key concept that aligned with the curriculum • Described student work • Explained what they were looking for in students' answers	• Key concept was vague or did not align with the curriculum • Did not describe what they were looking for in students' answers
Review/reflect	• Described trends in student work • Included evidence both of how students understood and where they still needed more help	• Only mentioned what students understood or what students missed • Provided very little detail about students' ideas
Adjust	• Provided rationale for a next step strategy that aligned with the students' ideas they previously described	• Did not provide rationale for using a particular next step strategy and/or • The next step strategy chosen was not appropriate to address the students' ideas they had previously described

manual said I should be looking for: erosion, deposition, the change of the shape of the pile" (Alicia, Interview 4). The teachers used curriculum-specific language to describe what concept they were targeting rather than state the concept in their own words or indicate that they understood the concept beyond what the curriculum defined.

For the third-grade teachers, the presence or absence of these particular terms often indicated whether or not the students understood the concept. For example, Sheri taught a lesson about properties of water on different surfaces in which students explored how surface tension creates a dome of water on a surface. After analyzing students' artifacts from the lesson, she recalled, "I looked at, Did they use the words 'surface tension'? and Can water can connect, hold itself together?" (Sheri, Interview 2). Although most students drew a dome of water on an impermeable surface, many did not also use the words surface tension in their explanation of the property of water that caused this observable phenomenon, which Sheri identified as evidence of incomplete conceptual understanding. Sheri only identified students' understanding as accurate and complete if they used the term surface tension in their answers (Sheri, Log 4). All three third-grade teachers reinforced this emphasis on terminology in the classroom, recording terms in public spaces in the classroom (Anna, Observation 4; Kim, Observation 3; Sheri, Observation 1) and asking the students to both repeat words verbally and record them in their notebooks (Anna, Observation 4; Sheri, Observation 1). In this way, the emphasis on terminology in the curriculum materials was reflected in teachers' formative assessment practices in the classroom.

Similar to the third-grade teachers, the three fifth-grade teachers (Alicia, Cathryn, and Matt) also mentioned looking for specific vocabulary words in evidence of students' thinking. There was evidence that they more frequently emphasized and supported students' use of appropriate vocabulary by explaining ideas behind specific terminology or curricular representations. Alicia, for example, mentioned that it was difficult to understand if students truly understood the concept behind specific terminology if they only used the word and didn't further expand on the idea, saying, "I would not count it as wrong because they understood what erosion meant, they just used different words to explain it but that it would be helpful to make sure that they added that word" (Alicia, Interview 4). Yet, although the fifth-grade teachers emphasized students' thinking underlying language use, they typically spoke about those ideas in general terms. In Matt's first lesson, for example, he noted he "was looking for answers . . . that demonstrated an understanding of what erosion and deposition are" (Matt, Log 1). In Matt's enacted lesson (Observation 1), he prompted students to articulate the definition of the term erosion as defined in the unit and used this definition as a core criterion in his analysis of students' artifacts from the lesson, This focus on curriculum-specific language, either for terminology or definitions, illustrates how teachers drew heavily on their curriculum materials to define Earth Science concepts rather than leveraging their own content knowledge to more broadly analyze and interpret artifacts of students' thinking.

Implementing Follow-up Instruction

When the teachers enacted follow-up instruction in response to their analysis of students' ideas, all six tended to use the same follow-up instructional strategies repeatedly. The choice of follow-up instructional strategies did not tend to focus on the particular ideas that the teachers identified in student work, but on using one or more of a select few

TABLE 4. Descriptive statistics for teachers' content exam and lesson log scores.

	M (*SD*)	Minimum Possible Score	Maximum Possible Score	Minimum Achieved Score	Maximum Achieved Score
Content exam score	8.82 (1.61)	0	12	5	12
Average log score	3.16 (0.48)	0	4	1.3	3.9
Anticipate	3.31 (0.92)	0	4	1	4
Review/reflect	3.22 (0.61)	0	4	1.25	4
Adjust	2.83 (0.82)	0	4	1	4

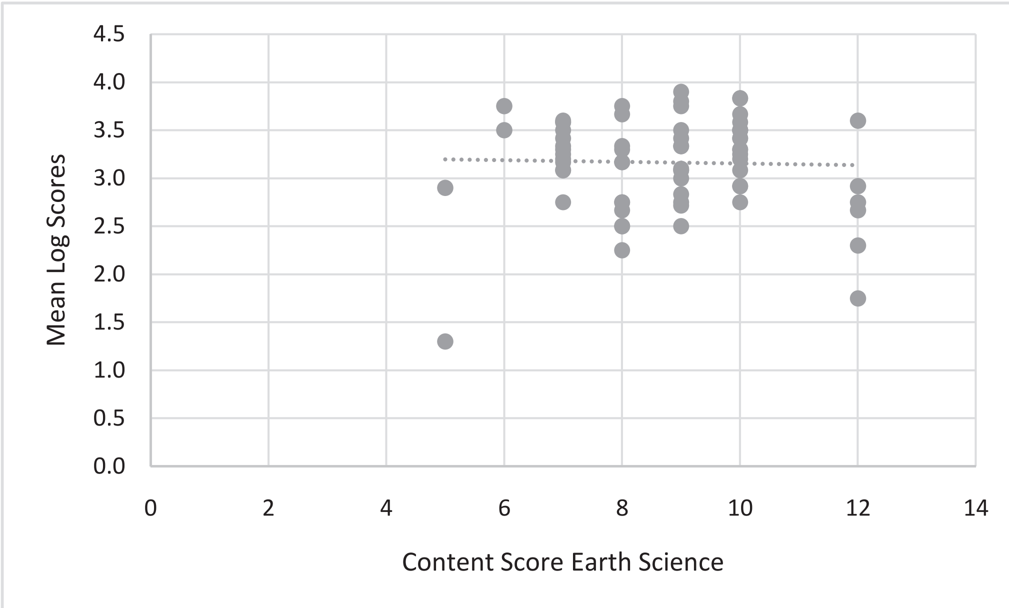

FIGURE 2. Mean log score compared to the content score for Earth Science items for each teacher.

instructional strategies with which they were already familiar. Their rationales for choosing these strategies revolved around the demands of learning a new strategy, both for the students and for themselves, and their level of comfort with the strategy itself. For example, Anna said, "The line of learning, we've done a lot. The color with revision [an instructional strategy], we've probably done those the most. . . . But definitely . . . you get comfortable doing [a particular] one" (Anna, Final interview). Anna's reflection was representative of all six teachers. In effect, the teachers' pedagogical reasoning was more frequently grounded in their own instructional priorities and needs than reflective of the learning needs of their students

A common characteristic of teachers' follow-up instruction was to have students share their ideas with other students to improve upon their responses. In follow-up instruction, teachers frequently chose examples of student work, often including samples they identified as exhibiting strong understanding of Earth Science concepts as defined by the units. They used the students' own work as a starting point for discussion about target concepts and they asked students to revise their own work after a discussion or reteaching of the concept (Alicia, Observation 3; Anna, Observation 4; Cathryn, Observation 2; Kim, Log 9; Matt, Observation 4; Sheri, Observation 2). Matt, for example, justified this approach in his second lesson, saying:

> *"I chose this strategy because I feel that by looking at classmate responses they might be able to help them understand the difference between erosion and deposition, and how they are related. Those students that need further understanding of these concepts will benefit from hearing classmates give constructive feedback to them."* (Matt, Log 2)

Here, Matt highlights the important link between the students' reflection on their own ideas, those articulated by their peers, and the key concept targeted by the curriculum.

However, in their follow-up instruction, the teachers frequently focused on an idea or component of broader concepts for which students may have possessed reasonable understanding, but simply did not write explanations or articulate their ideas using curriculum-defined terminology. Evidence from students' artifacts and classroom observations suggests that while these strategies may have supported students in articulating responses that aligned with representations of key concepts in the curriculum, they may have been less effective in furthering students' understanding of the content. For example, Sheri created a lesson in response to the student work described above in which students drew domes of water, but did not use the words *surface tension*. For the lesson, she brought in a metal bowl to give the students another example of how water forms a dome on metal surfaces (Sheri, Log 4 and Observation 2). Sheri focused on demonstrating water forming a dome on a metal surface again, a concept for which most of the students exhibited understanding of in the student artifacts from the previous lesson, rather than focusing on the problem she identified with students not using the correct terminology. By using these same strategies repeatedly, and by using examples of students' own work, the teachers' follow-up instruction most strongly resembled reteaching of key concepts from the Earth units as compared to targeted instructional support based upon identification of gaps in students' understanding of those concepts. While they may have provided new ways for students to consider the key concepts, these follow-up instructional approaches still typically relied on addressing the overall concept rather than on specific elements of those concepts for which students could benefit from further instruction.

Summary of Findings

This study began with an examination of teachers' Earth Science content knowledge in relation to their evaluation of students' ideas and selection of follow-up instructional strategies. However, quantitative analyses showed no significant relationship between teachers' Earth Science content knowledge and their engagement in formative assessment. Results from cross-case analysis of the six teachers confirm these findings. Rather than employ their knowledge of relevant Earth Science content, the teachers tended to ground their interpretation of students' thinking within a disciplinary frame defined by the Earth Science curriculum modules they used. When making decisions about follow-up instruction without the same level of discipline-specific, curriculum-embedded support, teachers tended to use the same follow-up instructional strategies regardless of the content or gaps they identified in students' thinking about water and earth materials.

SYNTHESIS AND DISCUSSION

Elementary students need opportunities to engage in scientific practices that are grounded in contemporary perspectives on science teaching and learning (Donovan and Bransford, 2005) and to learn about core disciplinary knowledge (NRC, 2007; NGSS Lead States, 2013). However, past studies have shown that they need explicit support in developing scientifically accurate explanations for Earth Science concepts, such as interactions between the hydrosphere and geosphere (Bar, 1989; Cheek, 2010; Forbes et al., 2015). Teachers can effectively scaffold students' learning by taking individual student progress into account, engaging and leveraging students' ideas, and cultivating student-centered science learning environments (Cowie and Bell, 1999; Bell and Cowie, 2001), all of which can lead to student learning gains (Ruiz-Primo and Furtak, 2006; Ayala et al., 2008; Shavelson et al., 2008). Prior studies have investigated preservice and inservice elementary teachers' use of formative assessment for science (Morrison and Lederman, 2003; Aschbacher and Alonzo, 2006; Otero and Nathan, 2008; Buck and Trauth-Nare, 2009; Coffey et al., 2011; Hammer et al., 2012; Morrison, 2013; Sabel, 2015). Findings from this study extend this body of previous research, providing empirical evidence for relationships between elementary teachers' knowledge of Earth Science concepts and formative assessment practices, how teachers use formative assessment to evaluate students' discipline-specific ideas, and the instructional reasoning they use to choose and implement follow-up instruction within the context of Earth Science instruction.

First, study findings suggest elementary teachers' Earth Science content knowledge does not influence their formative assessment practices. While variation was observed in teachers' disciplinary knowledge and formative assessment practices, higher levels of Earth Science content knowledge did not necessarily lead to more sophisticated formative assessment practices, or vice versa. Past research has shown that elementary teachers with stronger content knowledge are able to implement more effective teaching strategies, and have even indicated positive relationships between teachers' content knowledge and student learning outcomes (Hill et al., 2005; Falk, 2012), including for formative assessment (Aschbacher and Alonzo, 2006; Heritage et al., 2009). As a result of and in tandem with this body of empirical work, there have been increasing calls for attention to the disciplinary nature of formative assessment (Coffey et al., 2011; Gottheiner and Siegel, 2012), including in standards for teacher preparation (Brookhart, 2011).

Findings presented here suggest this relationship may be more complicated and multifaceted than often conceived. In particular, teachers may need support to learn how to connect Earth Science concepts to their pedagogical practices in the classroom. This missing connection could help explain findings from previous studies, which have shown that teachers do not understand how to use formative assessment or have sufficient content knowledge to do so (Otero and Nathan, 2008; Sabel et al., 2015). Although past work has shown that teachers with more content knowledge implement more effective teaching strategies (Hill et al., 2005; Falk, 2012), this may due to additional value-added supports for other instructional approaches, or greater understanding of instructional strategies. Therefore, while disciplinary knowledge remains important for teachers, findings here suggest that it alone is not sufficient for effective implementation of formative assessment without additional support for teachers to integrate what they know about the topic with what they do in the classroom. Thus, our finding extends existing research, presents counterevidence to conventional wisdom about the role of teachers' content knowledge in instructional practice, and strongly merits further study. They also highlight the need for novel programmatic models for preparing and supporting teachers to implement formative assessment.

Second, rather than draw upon their own Earth Science content knowledge, the elementary teachers in this study relied heavily on their Earth Science curriculum modules to interpret students' thinking. Specifically, they focused on the particular language and representations of content used in their instructional materials to describe target concepts for student learning and the evidence of understanding in students' responses. This finding could help explain the unexpected nonrelationship observed between teachers' Earth Science content knowledge and formative assessment practices. The instructional materials the teachers used were highly developed and emphasized particular key Earth Science concepts (FOSS, 2005a, b), which largely served as the lens through which teachers evaluated students' thinking. Whether or not teachers possessed strong content knowledge, the curricular resources largely dictated the disciplinary nature of both instruction and student experiences. This finding reinforces the importance of elementary science curriculum materials in shaping elementary teachers' science teaching practice (Biggers et al., 2013; Forbes et al., 2013; Zangori et al., 2013). Specifically, it highlights the importance of the curricular context, as well as the instructional resources themselves, in shaping how elementary teachers work towards goals for students' discipline-specific conceptual growth and interpret evidence of students' thinking. It also further supports our previous assertion that teachers need more support in learning how to leverage their own content knowledge of geoscience concepts in combination with their use of the curricular resources so that they can address issues with students' thinking that may be beyond the scope of misuse of vocabulary or misunderstandings of the curriculum-defined key concept.

Third, teachers' decisions about follow-up instruction typically did not effectively target gaps in student understanding about targeted Earth Science concepts. Both the RA strategy that teachers implemented and the curriculum modules they used provided the teachers with structure for instruction and organizing formative assessment in their classrooms (Kennedy et al., 2009). However, the follow-up instructional strategies that teachers used in this study were not a part of the curriculum, were not discipline-specific, and required the teachers to make decisions about how to implement appropriate strategies on their own. Teachers relied on the same instructional strategies largely independent of the conceptual context of student tasks. Although the teachers were able to determine that they should utilize follow-up instruction based on their evaluation of students' understanding, they did not typically choose follow-up instructional approaches based upon the learning needs of the students. Teachers may have difficulty in moving from the structure of the curriculum to instructional design, particularly if they do not leverage their own content knowledge and instructional decisions in ways that go beyond the curriculum. This finding joins other research in highlighting the need for disciplinary perspectives on formative assessment (Coffey et al., 2011), particularly follow-up instruction. It also reinforces previous research, showing that teachers tend to be more effective at identifying trends in students' thinking in relation to target concepts than making pedagogical decisions about subsequent instruction (Morrison and Lederman, 2003; Otero and Nathan, 2008; Buck and Trauth-Nare, 2009; Heritage et al., 2009; Hammer et al., 2012; Morrison, 2013; Sabel et al., 2015).

IMPLICATIONS AND CONCLUSION

Though engaging in formative assessment itself can serve as effective professional development for teachers (Ash and Levitt, 2003), findings from this study suggest that teachers may not engage their Earth Science content knowledge through instructional practice. These findings illustrate the need to provide teachers with explicit support in learning to evaluate students' ideas and select instructional strategies that address gaps in student understanding of core Earth Science concepts, such as interactions between the hydrosphere and geosphere. Teachers must develop the ability to diagnose students' thinking effectively, be empowered with tools that support their pedagogical decision making, and learn to draw upon a wide array of robust, discipline-specific instructional strategies to promote students' learning. Findings from this study inform the design of instruction about Earth systems, and therefore have important implications for teacher education for the geosciences (Marcum-Dietrich et al., 2011), both preservice and inservice, as well as for curriculum development.

First, these findings have important implications for teacher education. Teacher education and professional development should support teachers in developing an understanding the importance of using formative assessment, how to implement it in their classrooms, and the geoscience content and pedagogical knowledge to support the practice (Otero and Nathan, 2008; Buck and Trauth-Nare, 2009; Brookhart, 2011; Coffey et al., 2011). To address each of these outcomes effectively, experiences for teachers must foreground and emphasize the interrelatedness of Earth Science content knowledge and classroom teaching and learning. Formative assessment, which possesses discipline-specific elements and is grounded in theoretical perspectives on student-responsive science instruction, provides a framework for this integrated approach. Teachers can learn to activate content knowledge to evaluate students' ideas and implement innovative instructional approaches, design effective prompts to ensure alignment between student tasks and target concepts, unpack target concepts and identifying trends in artifacts that provide evidence of students' thinking, and select or design discipline-specific follow-up instruction that will align with the student ideas they discover. Experiences such as these, grounded in authentic classrooms settings, provide a crucial context for professional learning (Ash and Levitt, 2003). Teachers can learn disciplinary content through engagement in formative assessment, which can better prepare elementary teachers learn to support students' learning of Earth Science concepts effectively (Marcum-Dietrich et al., 2011). These are critical assumptions and design features of the professional development program emphasized here in year 2 and beyond.

Second, the curriculum materials that teachers use can also serve as crucial supports for teachers implementing formative assessment practices. As has been found in other studies (Biggers et al., 2013; Forbes et al., 2013; Zangori et al., 2013), the elementary teachers studied here relied heavily on their curricular resources for science by closely following the teacher's guide and using the terminology emphasized in the curriculum as a means of evaluating students' ideas. To support teachers more effectively in using formative assessment, geoscience curricular resources should support their elicitation and evaluation of students' thinking, as well as teachers' follow-up instruction. To support teachers' use of on-the-fly, planned-for, and curriculum-embedded types of formative assessment (Ayala et al., 2008; Shavelson et al., 2008), instructional materials could include suggested discussion probes to elicit students' ideas of Earth Science concepts, identify critical points in instructional sequences in which students may require particular scaffolding, and provide specific curriculum-aligned tasks that teachers can use as artifacts of students' thinking. To support their pedagogical reasoning about student-responsive instruction, they could identify follow-up instructional strategies that align with target concepts, planned instructional sequences, and research on students' learning about Earth Science topics. These kinds of supports require careful attention to the design of curriculum materials that provide teachers with tools, resources, and flexibility to support students to bridge gaps in their understanding. They also require that teachers have support to recognize when their curricular resources are inadequate for supporting some of the student thinking they may encounter in practice.

This mixed-methods study provides empirical evidence of elementary teachers' use of formative assessment to scaffold students' learning about interactions between water and earth materials. However, more research is needed to investigate relationships between teachers' knowledge of other disciplinary content in the geosciences and beyond (i.e., life and physical sciences) and their engagement in formative assessment practices. This study was also limited to teachers using only two different curriculum kits. Given the importance of curricular resources in elementary teachers' implementation of formative assessment practices, future research should examine how teachers implement formative assessment as part of the enactment of different programs focused on the same topics. Further, the present study should be replicated using other instruments to assess teachers' knowledge of disciplinary content to either validate or provide counter evidence to study findings. As part of ongoing work in this professional development program, we are investigating how teachers' formative assessment practices evolve over time and how their learning is optimally supported through ongoing professional development. Finally, all of these patterns in teachers' formative assessment practices should be studied in relation to discipline-specific student learning outcomes.

Acknowledgments

Cory Forbes and Jaime Sabel are shared first authors and appear in alphabetical order. Mandy Biggers is second author. This work is funded by the Spencer Foundation, the Iowa Board of Regents, and the Title IIA Improving Teacher Quality: State Grants Program. However, any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors. We appreciate the interest and cooperation of Leslie Flynn, Christopher Soldat, Jeanne Bancroft, Kathy Long, Madison Fontana, and the elementary teachers who made this research possible. We also thank Laura Zangori, Tina Vo, and anonymous reviewers for their help in thinking about these issues and their thoughtful comments on earlier versions of this paper.

References

Aschbacher, P., and Alonzo, A. 2006. Examining the utility of elementary science notebooks for formative assessment purposes. *Educational Assessment,* 11:179–203.

Ash, D., and Levitt, K. 2003. Working within the zone of proximal development: Formative assessment as professional development. *Journal of Science Teacher Education,* 14:23–48.

Ayala, C.C., Shavelson, R.J., Ruiz-Primo, M.A., Brandon, P.R., Yin, Y., Furtak, E.M., Young, D.B., and Tomita, M.K. 2008. From formal embedded assessments to reflective lessons: The development of formative assessment studies. *Applied Measurement in Education,* 21:315–334.

Ball, D.L., and Forzani, F.M. 2009. The work of teaching and the challenge for teacher education. *Journal of Teacher Education,* 60:497–511.

Bar, V. 1989. Childrens' views about the water cycle. *Science Education,* 73:481–500.

Bell, B., and Cowie, B. 2001. The characteristic of formative assessment in science education. *Science Education,* 85:536–553.

Biggers, M., Forbes, C.T., and Zangori, L. 2013. Elementary teachers' curriculum design and pedagogical reasoning for supporting students' comparison and evaluation of evidence-based explanations. *Elementary School Journal,* 114(1):48–72.

Brookhart, S.M. 2011. Educational assessment knowledge and skills for teachers. *Educational Measurement: Issues and Practice,* 30:3–12.

Buck, G.A., and Trauth-Nare, A.E. 2009. Preparing teachers to make the formative assessment process integral to science teaching and learning. *Journal of Science Teacher Education,* 20:475–494.

Camburn, E., and Barnes, C.A. 2004. Assessing the validity of a language arts instruction log through triangulation. *Elementary School Journal,* 105:49–73.

Cheek, K.A. 2010. Commentary: A summary and analysis of twenty-seven years of geoscience conceptions research. *Journal of Geoscience Education,* 58:122–134.

Clark, L.A., and Watson, D. 1995. Constructing validity: Basic issues in objective scale development. *Psychological Assessment,* 7:309–319.

Coffey, J.E., Hammer, D., Levin, D.M., and Grant, T. 2011. The missing disciplinary substance of formative assessment. *Journal of Research in Science Teaching,* 48:1109–1136.

Cowie, B., and Bell, B. 1999. A model of formative assessment in science education. *Assessment in Education,* 6:101–116.

Cresswell, J.W., and Plano Clark V.L. 2011. Designing and conducting mixed methods research. Los Angeles, CA: Sage.

Donovan, M.S., and Bransford, J.D., eds. 2005. How students learn: Science in the classroom. Washington, DC: National Academies Press.

Falk, A. (2011). Teachers learning from professional development in elementary science: Reciprocal relations between formative assessment and pedagogical content knowledge. *Science Education,* 96:265–290.

Forbes, C.T., Biggers, M., and Zangori, L. 2013. Investigating essential characteristics of scientific practices in elementary science learning environments. *School Science and Mathematics,* 113(4):180–190.

Forbes, C.T., Zangori, L., and Schwarz, C.V. 2015. Empirical validation of integrated learning performances for hydrologic phenomena: 3rd-grade students' model-driven explanation-construction. *Journal of Research in Science Teaching,* (early view).

Full Option Science System (FOSS). 2005a. FOSS Teacher Guide: Water. Berkeley, CA: Delta Education, Inc.

Full Option Science System (FOSS). 2005b. FOSS Teacher Guide: Landforms. Berkeley, CA: Delta Education, Inc.

Gottheiner, D.M., and Siegel, M.A. 2012. Experienced middle school science teachers' assessment literacy: Investigating knowledge of students' conceptions in genetics and ways to shape instruction. *Journal of Science Teacher Education,* 23:531–557.

Hammer, D., Goldberg, F., and Fargason, S. 2012. Responsive teaching and the beginnings of energy in a third grade classroom. *Review of Science, Mathematics, and ICT Education,* 6:51–72.

Heller, J.I., Daehler, K.R., Wong, N., Shinohara, M., and Miratrix, L.W. 2012. Differential effects of three professional development models on teacher knowledge and student achievement in elementary science. *Journal of Research in Science Teaching,* 49:333–362.

Heritage, M., Kim, J., Vendlinski, T., and Herman, J. 2009. From evidence to action: A seamless process in formative assessment? *Educational Measurement: Issues and Practice,* 28:24–31.

Hill, H.C., Rowan, B., and Ball, D.L. 2005. Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal,* 42:371–406.

Kennedy, C., Long, K., and Camins, A. 2009. The reflective assessment technique. *Science and Children,* 47:50–34.

Littell, R.C., Milliken, G.A., Stroup, W.W., Wolfinger, R.D., and Schabenberger, O. 2006. SAS for mixed models, 2nd ed. Cary, NC: SAS Institute, Inc.

Marcum-Dietrich, N., Marquez, L., Gill, S.E., and Medved, C. 2011. No teacher left inside: Preparing a new generation of teachers. *Journal of Geoscience Education,* 59:1–4.

Merriam, S.B. 2009. Qualitative research: A guide to design and implementation. San Francisco, CA: Jossey-Bass.

Miles, M.B., and Huberman, A.M. 1994. Qualitative data analysis. Newbury Park, CA: Sage.

Morrison, J.A. 2013. Exploring exemplary elementary teachers' conceptions and implementation of inquiry science. *Journal of Science Teacher Education,* 24:573–588.

Morrison, J.A., and Lederman, N.G. 2003. Science teachers' diagnosis and understanding of students' preconceptions. *Science Education,* 87:849–867.

National Research Council (NRC). 2007. Taking science to school: Learning and teaching science in grades K–8. Washington, DC: The National Academies Press.

Next Generation Science Standards (NGSS) Lead States. 2013.

Next generation science standards. For states, by states. Washington, DC: The National Academies Press.

Penuel, W.R., McWilliams, H., McAuliffe, C., Benbow, A.E., Mably, C., and Hayden, M.M. 2009. Teaching for understanding in Earth Science: Comparing impacts on planning and instruction in three professional development designs for middle school science teachers. *Journal of Science Teacher Education,* 20:415–436.

Otero, V.K., and Nathan, M.J. 2008. Preservice elementary teachers' views of their students' prior knowledge. *Journal of Research in Science Teaching,* 45:497–523.

Rowan, B., and Correnti, R. 2009. Studying reading instruction with teacher logs: Lessons from the study of instructional improvement. *Educational Researcher,* 38:120–131.

Ruiz-Primo, M.A., and Furtak, E.M. 2006. Informal formative assessment and scientific inquiry: Exploring teachers' practices and student learning. *Educational Assessment,* 11:205–235.

Sabel, J.L., Forbes, C.T., and Zangori, L. 2015. Content knowledge and formative assessment integration in a life sciences methods course for preservice teachers. *Journal of Science Teacher Education,* 26(4):419–445.

Sadler, P.M., Coyle, H., Cook Smith, N., Miller, J., Mintzes, J., Tanner, K., and Murray, J. 2013. Assessing the life science knowledge of students and teachers represented by the K–8 National Science Standards. *CBE—Life Sciences Education,* 12: 553–575.

Sadler, P.M., Coyle, H., Miller, J.L., Cook-Smith, N., Dussault, M., and Gould, R.R. 2010. The astronomy and space science concept inventory: Development and validation of assessment instruments aligns with the K–12 National Science Standards. *Astronomy Education Review,* 8(1):010111.

Shavelson, R.J., Young, D.B., Ayala, C.C., Brandon, P.R., Furtak, E.M., Ruiz-Primo, M.A., Tomita, M.K., and Yin, Y. 2008. On the impact of curriculum-embedded formative assessment on learning: A collaboration between curriculum and assessment developers. *Applied Measurement in Education,* 21:295–314.

Wendt, J.L., and Rockinson-Szapkiw, A. 2014. The effect of online collaboration on middle school student science misconceptions as an aspect of science literacy. *Journal of Research in Science Teaching,* 51:1103–1118.

Yin, R.K. 2009. Case study research: Design and methods. Thousand Oaks, CA: Sage.

Zangori, L., Forbes, C.T., and Biggers, M. 2013. Fostering student sense making in elementary science learning environments: Elementary teachers' use of science curriculum materials to promote explanation-construction. *Journal of Research in Science Teaching,* 50(8), 887–1017.