

Fall 1-8-2016

Keeley Probes as a Tool for Uncovering Student Ideas: How Do Teachers Use Formative Assessment Probes to Plan and Adapt Instruction?

Kalin Tobler
Portland State University

Follow this and additional works at: https://pdxscholar.library.pdx.edu/open_access_etds



Part of the [Science and Mathematics Education Commons](#)

Let us know how access to this document benefits you.

Recommended Citation

Tobler, Kalin, "Keeley Probes as a Tool for Uncovering Student Ideas: How Do Teachers Use Formative Assessment Probes to Plan and Adapt Instruction?" (2016). *Dissertations and Theses*. Paper 2650. <https://doi.org/10.15760/etd.2646>

This Thesis is brought to you for free and open access. It has been accepted for inclusion in Dissertations and Theses by an authorized administrator of PDXScholar. Please contact us if we can make this document more accessible: pdxscholar@pdx.edu.

Keeley Probes as a Tool for Uncovering Student Ideas: How Do Teachers Use
Formative Assessment Probes to Plan and Adapt Instruction?

by

Kalin Tobler

A thesis submitted in partial fulfillment of the
requirements for the degree of

Master of Science Teaching
in
General Science

Thesis Committee:
William Becker, Chair
Cary Sneider
Stephanie Wagner

Portland State University
2015

Abstract

Formative assessment probes, known as Keeley probes, are one tool teachers use to reveal students' scientific misconceptions, so that they can move them closer to conceptual understanding. The purpose of this research was to document how four elementary school teachers used formative assessment probes to plan and adapt instruction to improve student learning. Specifically:

- How did teachers choose appropriate probes? What learning goals did teachers hope to address by using the probe??
- What instructional sequences did teachers envision when planning to use a probe?
- What did teachers notice when analyzing student data from a probe?
- How did teachers use the information to modify their instructional practice?

This exploratory study addresses key issues by exploring through qualitative methods how four elementary teachers used Keeley formative assessment probes in the classroom through a series of individual and group interviews. The results, reported as case studies and themes, indicate that Keeley probes may be used to help teachers strengthen their pedagogical content knowledge and as an anchor for classroom discussions. Teachers reported that students were highly engaged when considering Keeley probing questions. Teachers in this study had questions about how to analyze data collected through formative assessment, and what instructional steps they needed to take to address misconceptions.

The central finding of the study is that a teacher's subject-area knowledge as well as the ability to identify students' misconceptions and make instructional decisions based on those ideas, both elements of pedagogical content knowledge, play a key role in how effectively teachers use Keeley formative assessment probes towards improving learning.

Ultimately, this study showed that while the use of Keeley probes did improve opportunities for students to deepen scientific understanding, a gap still exists between the potential of formative assessment and the practical work of integrating ongoing formative assessment to improve teaching and learning.

This exploratory study underlines the need for a new approach in professional development for elementary science teachers, and sheds light on what happens when teachers try Keeley probes, a promising formative assessment tool and strategy, in the real world of the classroom.

Acknowledgements

I would like to thank the following individuals for their support of this Master's thesis study. Stephanie Wagner, MST Program Coordinator, invited my participation as a trainer and instructional coach for NGSS Workshops, connected me with the teachers and principal at the site of my research, and remained flexible and sympathetic throughout my Master's degree program. The teachers and principal at the beautiful school where I conducted my study participated thoughtfully and sincerely in research activities and inspired me as a teacher-leader. Little Vikings, the flexible childcare at Portland State University, allowed me to focus on research and learning knowing my young daughter was well-cared-for. Cary Sneider, Associate Professor in the Center for Science Education at Portland State University, provided valuable support and encouragement throughout my thesis research activities and my entire Master's degree program. Kalin Tobler

Table of Contents

Abstract.....	i
Acknowledgements.....	iii
List of Tables	vi
Chapter 1 Introduction	1
Chapter 2 Review of Literature.....	7
Chapter 3 Methods.....	31
Chapter 4 Results	45
Chapter 5 Discussion	75
References.....	89

Appendices

A. “Respiration” Probe.....	92
B. “Mountaintop Fossil” Probe.....	93
C. “Is It Matter?” Probe	94
D. Sample Teacher Responses For Each Conceptual Code.....	95
E. NGSS Correlations: “Respiration.....	97
F. NGSS Correlations: “Mountaintop Fossil.....	98
G. NGSS Correlations: “Is It Matter	99
H. Sample Student Responses: “Respiration	101
I. Sample Student Responses: “Mountaintop Fossil.....	103

J. Jason’s Sample Student Responses: “Is It Matter.....104

K. Christopher’s Sample Student Responses: “Is It Matter.....105

List of Tables

Table 1. Description of Three Formative Assessment Probes	38
Table 2. Research Elements and Timeline	41
Table 3. Number of Students Who Marked Organism as Respiring	51
Table 4. Number of Jason's Students Who Marked Substance as Matter	55
Table 5. Number of Christopher' Students Who Marked Substance as Matter.....	59

Introduction

More and more teachers and educational leaders are looking at the way assessment can play a role in enhancing student learning instead of just measuring it at the end of a unit or school year. Formative assessment, or assessment for learning, prioritizes collecting data about students' scientific ideas before and during instruction, and then using that information to improve teaching. Black and his colleagues (2004) give this definition of formative assessment:

Assessment for learning is any assessment for which the first priority in its design and practice is to serve the purpose of promoting pupils' learning. It thus differs from assessment designed primarily to serve the purposes of accountability, or of ranking, or of certifying competence. An assessment activity can help learning if it provides information to be used as feedback, by teachers, and by their pupils, in assessing themselves and each other, to modify the teaching and learning activities in which they are engaged (Black, et al, 2004, p. 11).

Compelling evidence supports the claim that formative assessment has a positive impact on student achievement, especially for low-performing students (Black & Wiliam, 1998a; Furtak & Ruiz-Primo, 2008; Sadler, et al., 2013). In fact, formative assessment has been proclaimed essential by a variety of state and national science organizations, universities, school districts, and assessment specialists (NRC, 2012; Orland & Anderson, 2013). Unfortunately, there is still a lack of focus on the real work of teaching and learning in the classroom (Black & Wiliam, 1998b).

The Next Generation Science Standards (NGSS) are built on the notion that learning is a developmental progression in which children continually build on and revise their knowledge and abilities, starting from their observations and initial conceptions

about how the world works. “The goal is to guide their knowledge toward a more scientifically based and coherent view of the natural sciences and engineering, as well as of the ways in which they are pursued and their results can be used” (NRC, 2012, p. 11). By embedding formative assessment throughout the lesson cycle, teachers may be more likely to meet this goal.

Transforming the curriculum to incorporate formative assessment is not simple. Wylie and Heritage (2010) assert that embracing and implementing formative assessment means huge changes for most teachers—changes in their views of themselves as teachers and in their understanding of the relationship between instruction and assessment. According to Heritage and her colleagues, teachers have difficulty using information gathered through formative assessment to inform the next instructional steps they will take (Heritage, et al. 2009). Teachers, professional developers, and educational leaders need to know more about how teachers and instruction are transformed through the implementation of a variety of formative assessment strategies and tools.

To be effective, teachers need subject matter (content) knowledge as well as knowledge of instructional approaches and methods (pedagogical knowledge). In 1986, Schulman introduced the idea of pedagogical content knowledge (PCK), “the blending of content and pedagogy into an understanding of how particular topics, problems, or issues are organized, represented, and adapted to the diverse interests and abilities of learners, and presented for instruction,” (Schulman, 1987, p.8) as the knowledge base for teaching. A science teacher’s knowledge, used to help students understand specific concepts, is different from a scientist’s knowledge. Cochran, DeRuiter, & King (1993) described the

integration of four major components in a model of pedagogical content knowledge: 1) subject matter knowledge, 2) pedagogical knowledge, 3) teacher's knowledge of students' abilities and learning strategies, ages and developmental levels, attitudes, motivations, and prior knowledge of the concepts being taught; and 4) teachers' understanding of the social, political, cultural and physical environments in which students are asked to learn.

In science particularly, teachers need to learn about their students' misconceptions in order to move them toward deeper conceptual understanding. Students come to the science classroom with many ideas that differ from accepted scientific concepts, and research shows that if new information does not fit with students' prior experiences and established patterns of thinking, students simply modify the new information to fit in with what they already thought, instead of changing their conceptual framework (Gooding & Metz, 2011).

The National Research Council (NRC) suggests that to break down student misconceptions, teachers need to identify those misconceptions, support students to confront their own ideas, and then provide learning experiences to reconstruct and internalize their new knowledge (NRC 1997). Sadler and his colleagues found that teachers who have subject matter knowledge and are able to identify common student misconceptions surrounding each particular concept help students learn more (Sadler, et al., 2013). Ample research shows common misconceptions about a variety of science topics (AAAS, 1993; Driver, 1994;), but each individual student holds unique, often deeply-rooted conceptions about the natural world. Therefore, science teachers must be

able to identify each student's misconceptions. This is why formative assessment is so important in science classrooms.

Eliciting evidence of students' ideas and learning is one of five key strategies proposed by Wiliam and Thompson (2007). Establishing where learners are, where they are going, and how to get there provide the framework for formative assessment. Without effective formative assessment tools and strategies, teachers proceed through a lesson or curriculum with little knowledge about what their students are actually thinking or learning.

If the substantial rewards promised by research about formative assessment are to be realized, each teacher must find a way to incorporate this approach in the classroom, but developing quality assessments is challenging and time-intensive. To help support the work of teachers and teacher-leaders eager to transform their teaching, Page Keeley, Senior Science Program Director at the Maine Mathematics and Science Alliance, has led the effort to develop effective, ready-made formative assessment tools that teachers can use to elicit students' ideas and inform instructional practices by designing probing questions that reveal students' preconceptions. Keeley calls the formative assessment questions and strategies "probes". The National Science Teachers Association now publishes a series called *Uncovering Student Ideas*, which started with the first book, *Vol. 1: Uncovering Student Ideas: 25 Formative Assessment Probes* (Keeley, 2005).

By definition, formative assessment is an approach, not one specific strategy. Probing questions like those developed and field-tested by Keeley's teams are flexible

and incorporate a variety of what Keeley calls Formative Assessment Classroom Techniques (FACTs, Keeley, 2008). Each probe is designed to target one or more scientific concepts, and provides an example of an effective instructional strategy or activity that can be used to develop understanding of the ideas targeted by the probe. Probes can be used before instruction, during a lesson, or after a lesson; teachers may analyze the written responses of students to find out the variety of ideas held, prepare for a scientific investigation, or may use them to prompt classroom discussion.

As with any tool or strategy for formative assessment, it is essential that the data collected about student ideas be used to adapt instruction so that students have the opportunity to revise their thinking and build a more accurate conceptual framework. While numerous anecdotal accounts are referenced in the Uncovering Student Ideas series (Keeley, Eberle, & Farrin, 2005), little research explores the ways teachers actually use the probes in the classroom. The purpose of this research is to document the different ways four elementary teachers use the Keeley probes to elicit student ideas and modify teaching. What are the ways that teachers use Keeley formative assessment probes to learn about their students' pre-conceptions, and how do they use that information to adapt their instruction? Specifically:

- How did teachers choose appropriate probes? What learning goals did teachers hope to address by using the probe?
- What instructional sequences did teachers envision when planning to use a probe?
- What did teachers notice when analyzing student data from a probe?
- How did teachers use the information to modify their instructional practice?

Data was collected via interviews with teachers and from students' written responses. Four teachers were interviewed three times—once before the lesson in which

they planned to use the Keeley probe, once immediately after the lesson, and once after he or she has had a chance to address the students' ideas discovered through analysis of the Keeley probe. To nurture dialogue and allow for discourse on possible themes that emerge, the final interview was a group interview. I offer a theoretical framework using conceptual categories which arose throughout teacher interviews.

This study revealed how a small group of experienced, motivated teachers attempt to translate the idea of formative assessment into everyday practice. The data provides a living example of how teachers use one formative assessment tool, the Keeley probe, to bring to light what students are thinking and reshape their instruction accordingly. By illuminating the important details about the process teachers engaged in to detect and then make instructional decisions to address students' misconceptions, the study provides valuable information which should be considered in designing assessment courses for student teachers and for professional development. Ultimately, this study contributes to the scholarship needed to realize the tremendous potential of formative assessment to improve student learning.

Literature Review

This following literature review includes articles and reports that address five related topics: 1) The Power of Formative Assessment to Improve Learning: Seminal Studies and Continuing Research; 2) Critical Reviews of Prominent Research; 3) Blending Content and Pedagogy: Pedagogical Content Knowledge; 4) The Importance of Uncovering Student Ideas: Taking Action to Improve Instruction; 5) Living Examples of Implementation: Strategies, Approaches, and Impact on Student Learning.

The Power of Formative Assessment: Seminal Studies and Continuing Research: Black and Wiliam

In 1998, Paul Black and Dylan Wiliam published a seminal piece on formative assessment, which is frequently cited as clear and unequivocal evidence that formative assessment improves student achievement. In this well-known meta-analysis, Black and Wiliam (1998a) analyzed more than 500 research studies including their own research, to answer three questions:

- . 1) Is there evidence that improving formative assessment raises standards?
- . 2) Is there evidence that there is room for improvement?
- . 3) Is there evidence about how to improve formative assessment?

The authors concluded,

All these studies show that innovations that include strengthening the practice of formative assessment produce significant and often substantial learning gains. These studies range over age groups from 5-year-olds to university

undergraduates, across several school subjects and over several countries (Black & William, 1998b, p. 3).

Black and William also provided evidence that formative assessment had the potential to help close the achievement gap, while raising student achievement for all learners (Black & William, 1998a).

The authors provided this definition for formative assessment: We use the general term assessment to refer to all those activities undertaken by teachers—and by their students in assessing themselves—that provide information to be used as feedback to modify teaching and learning activities. Such assessment becomes formative when the evidence is actually used to adapt the teaching to meet student needs” (Black & William, 1998b, p. 2).

The authors described these essential components of effective formative assessment:

- specific feedback to students about the quality of work and what can be done to improve;
- student self-assessment
- students’ active involvement in understanding learning goals and targets
- instructional adjustments that take into account the results of assessment
- recognition of the profound impact assessment has on the motivation and self-esteem of students
- productive peer evaluation

Follow-Up Projects.

Black and William continued their inquiry into formative assessment with a number of follow-up projects. In one project, the research team explored how 24 teachers in Oxfordshire and Medway, England, used formative assessment strategies in the classroom (Black, Harrison, Lee, Marshall, & William, 2004). Each teacher chose how

they would improve their use of formative assessment, and decided what the “output” measure would be for his or her class—the national school-leaving examination, national tests, or end-of-module-test scores. The researchers set up a mini-experiment for each teacher and identified a comparison class, often a class taught by the same teacher in previous years to highlight any possible achievement gains made from the increased use of formative assessment. The authors found significant achievement gains for students of 19 of the teachers on whom the research team had complete data.

The research team aggregated the results by calculating the “standardized effect size,” finding the average effect size was around 0.3 standard deviations. “Such improvements, produced across a school, would raise a school in the lower quartile of the national performance tables to well above average” (Black et al, 2004, p. 11).

Teachers implemented strategies from four categories:

- 1) questioning: teachers planned questions and allowed appropriate wait time.
- 2) feedback through grading: teachers gave feedback that was designed to cause thinking.
- 3) peer- and self-assessment: teachers ensured that students were given enough time during lessons to evaluate their own work and that of others.
- 4) the formative use of summative tests.

Black and his colleagues (2004) also listened to teachers’ insights regarding the impact of his or her efforts to improve their use of the chosen strategy. Two excerpts from the teacher interviews demonstrated what happened when teachers began to listen more attentively to student ideas: “My whole teaching style has become more interactive. Instead of showing how to find solutions, a question is asked and pupils are

given time to explore answers together” – Nancy, Riverside School (Black, et al, 2004, p. 12).

Another teacher in the study went on to say,

There was a definite transition at some point, from focusing on what I was putting into the process, to what the students were contributing. It became obvious that one way to make significant sustainable change was to get the students doing more of the thinking. I then began to search for ways to make the learning process more transparent to the students. Indeed, I now spend my time looking for ways to get students to take responsibility for their learning and at the same time making the learning more collaborative.” –Tom (Black et al, 2004, p. 19)

The work of Black and Wiliam has shaped national and international conversations about the role of assessment in student learning, informed policy briefs (Orland & Anderson, 2013) and helped set the research agenda on assessment for the educational community.

Critical Reviews: A Call for Better Methodology and More Focused Attention on Student Ideas and Reasoning Within a Discipline

Although the value of formative assessment is widely accepted in the field of education, critical reviews by a variety of researchers made the case for improving methodology, cognitive-domain specificity, and focusing more attention on student ideas instead of strategies.

Dunn and Mulvenon (Dunn & Mulvenon, 2009) reviewed prominent literature on formative assessment, including the seminal Black & William study (1998a), concluding that there is a need to clarify terminology, produce empirical evidence supporting the impact of formative assessment on academic achievement, and establish a sound

research-validated framework for best practices in formative assessment. The authors examined each study cited by Black & William to support their claim that formative assessment has a positive impact on student performance. Dunn and Mulvenon claimed that concerns about generalizability, validity, effect sizes, methodological problems, small sample sizes, and lack of accounting for other factors such as teacher effects render Black & William's study inconclusive. While the authors acknowledged that the research discussed "does provide some support for the impact of formative assessment on student achievement" (p. 9), they highlighted the need for common terminology and better research design to evaluate the impact of formative assessment.

In another critical review, Bennett (2011) examined prominent research and writings about formative assessment. His purpose was to call for a more meaningful definition of formative assessment, recognize the inferential nature of assessment and emphasize the importance of conceptualizing formative assessment within specific domains.

Bennett examined six issues of concern regarding formative assessment:

- . 1) definition of formative assessment
- . 2) effectiveness claims
- . 3) domain considerations
- . 4) the need for more attention to educational measurement principles
- . 5) issues of professional development
- . 6) the impact of the educational system on the effectiveness and implementation of formative assessment.

Definition. After tracing the origins of the idea of formative assessment, Bennett described how formative assessment differs from summative assessment in purpose: the results of formative assessment are used for modification and improvement, while summative assessment judges achievement. Bennett (2011) quotes Bloom in clarifying

the purpose of formative assessment, "...to provide feedback and correctives at each stage in the teaching-learning process" (as cited on p. 6).

According to Bennett, there are two general schools of thought about the definition of formative assessment. One side believes formative assessment refers to an instrument, as in a diagnostic test, an interim assessment or an item bank, which would produce diagnostic scores. This view is common among test publishers.

The other side of the split holds the view that, "Formative assessment is a process used by teachers and students during instruction that provides feedback to adjust ongoing teaching and learning to improve students' achievement of intended outcomes" (Bennett, 2011, p. 6). In other words, formative assessment is a process, not an instrument. This view is more common among educators and researchers. Emphasis is on the actual use of results to adapt teaching to meet student needs.

Effectiveness claims. Bennett interprets effectiveness research such as the Black & William study (1998) to suggest that general practices associated with formative assessment can facilitate learning, but benefits vary widely in kind and size and from one subpopulation of students to the next. He calls for researchers to be more responsible in efficacy claims, and for educators to look for evidence of the benefits students will enjoy in the particular context in which they teach before jumping onto the formative assessment bandwagon.

Domain dependency. Bennett brings up two issues regarding domain: 1) the need for teachers to have deep cognitive-domain knowledge (physics knowledge and practices

differ significantly from life science, for example) in order to be able to know what questions to ask students and what actions to take to adjust instruction, and 2) “That deep cognitive-domain understanding includes the processes, strategies and knowledge important for proficiency in a domain, the habits of mind that characterize the community of practice in that domain, and the features of tasks that engage those elements” (p. 15). He argued that formative assessment should be embedded within the curriculum to address the specific understandings and practices required for deep understanding within a discipline.

Measurement. Bennett defined educational measurement as involving four activities: designing opportunities to gather evidence, collecting evidence, interpreting it, and acting on interpretations. Since formative assessment is an inferential process—educators make guesses about what students actually understand—it is difficult to be sure if conjecture about what adjustments to make to the curriculum is accurate. He also pointed out that formative inferences might be influenced by gender, race, ethnicity, disability, English language proficiency, or other student characteristics. To address this inferential uncertainty, he recommended considering data from multiple sources and grounding action in a “sound cognitive-domain model” (Bennett, 2011, p. 18).

Professional development. Bennett argued that teachers need substantial knowledge and time to implement the most effective formative assessment, deepen domain understanding, and to reflect upon their own experiences with the process and tools used in formative assessment. Bennett called for engagement in “iterative cycles of

use, reflection, adaptation, and even creation,” (Bennett, 2011, p. 19) to help teachers integrate deep domain understanding with the methodology of formative assessment.

The system. According to Bennett, the effectiveness of formative assessment is limited by the nature of the larger system in which it is embedded. In order to function effectively, formative and summative assessments need to be aligned, and teachers must acquire the skills needed to use assessment. The bigger challenge he identified was the need to change the system of multiple-choice summative assessments in order to have maximum impact on learning and instruction. Although Bennett recognized the potential of formative assessment to transform classroom instruction and learning, his critique described numerous factors that need to be considered to realize the maximum benefit from formative assessment.

In another critique, Coffey, et al. (2011) called for a focus not on strategies for the teacher, but on, “...becoming more aware of and responsive to student thinking, without the benefit of any particular strategies” (Coffey, Hammer, Levin, & Grant, 2011, p. 22).

In *The Missing Disciplinary Substance of Formative Assessment*, Coffey, et al. (2011) re-examined prominent research on formative assessment to support their claim that researchers have not been paying attention to the most important aspect of formative assessment—the substance of student ideas and reasoning, and the guidance teachers could provide towards disciplinary understanding. According to the researchers, “Formative assessment should be understood and presented as nothing other than genuine

engagement with ideas, which includes being responsive to students' ideas and using them to inform next moves" (p. 1129).

Coffey and her colleagues (2011) analyzed four transcripts of classroom discussions from prominent research to argue that the literature on formative assessment overlooked the substance of what teachers (and students) should be assessing. They chose these particular examples to argue the need to shift thinking towards deeper awareness and responsiveness to students' ideas and away from focusing on target vocabulary or the "correct answer". The authors supported three interrelated claims:

1. There is little discussion about the substance of student thinking.
2. There is a tacit presumption of "content" as a body of correct information, centered on terminology and selected in advance of lesson objectives.
3. Assessment is discussed in terms of particular strategies, techniques, and procedures, distinct from other teaching and learning activities.

In the first excerpt, (from Black, Harrison, Lee, Marshall, & William, 2003), the authors argued that, contrary to the claims of Black and his colleagues, there was no evidence the teacher tried to explore student understanding or unpack student ideas. According to Coffey, et al. (2011), the evidence suggested the teacher was more focused on steering students toward specific target knowledge about photosynthesis. In the second example, a high school discussion of density, Coffey et al. critiqued the claim of researchers that the teacher was engaging students with questions to find out what they understood, arguing that the teacher did not focus on the ideas for their own value, but rather, was looking for the correct answer (Coffey et al., 2011).

The authors then provided examples of teachers that were more attentive to student ideas and reasoning. Terry, teaching a ninth grade biology class, intended to present a 15- minute review of matter, atoms, and molecules as part of a unit on the chemistry of life. By listening to student responses and posing questions, Terry noticed students had little understanding of the material he thought would only need a quick review. Terry changed his goals for the lesson in response to what he heard, guided students to explore the concept of air as matter through an impromptu scientific investigation, and then planned a new learning activity the next day. This classroom discussion showed evidence of moving students toward both conceptual understanding and, more importantly, using reasoning and observations to engage in scientific debate without the use of particular formative assessment strategies.

All three critiques recognized the immense potential for formative assessment to transform teaching and learning. Research published by both Bennett (2011) and Dunn and Mulvenon (2009) called for better research methodology to measure the impact of formative assessment on student learning.

Blending Content and Pedagogy: Pedagogical Content Knowledge

In 1987, Lee Schulman supported the call for large-scale reforms to improve teaching by introducing the notion of pedagogical content knowledge (PCK). In his widely cited article, *Knowledge and teaching: Foundations of the new reform (1987)*, Schulman described how PCK blends the content knowledge a teacher is expected to master in a particular subject area with the strategies and knowledge of how that information may best be represented in order for students to understand:

But the key to distinguishing the knowledge base of teaching lies at the intersection of content and pedagogy, in the capacity of a teacher to transform the content knowledge he or she possesses into forms that are pedagogically powerful and yet adaptive to the variations in ability and background present by students (Schulman, 1987, p. 15).

Therefore, it is not enough for a teacher to understand the subject or to have deep pedagogical knowledge—teachers need both content area knowledge and pedagogical reasoning.

Schulman presented his model of pedagogical reasoning (Schulman, 1987) as a framework for a knowledge base for teaching which articulates how particular kinds of content knowledge and pedagogical strategies combine to make effective teachers. He emphasized the importance of preparing teachers more comprehensively in order to foster individual excellence as well as providing equality of opportunity and equity among students of different backgrounds and cultures.

Schulman identified the first element of PCK as comprehension. According to Schulman, “To teach is first to understand. We ask that the teacher comprehend critically a set of ideas to be taught. We expect teachers to understand what they teach and, when possible, to understand it in several ways,” Schulman stated. “They should understand how a given idea relates to other ideas within the same subject area and to ideas in other subjects as well” (p. 14).

Next, transformation must occur. Teachers must be able to move from their own personal comprehension of the subject matter to preparation for the comprehension of others. Schulman listed a variety of processes that are required to transform ideas:

- preparation (critical interpretation—structuring and segmenting subject matter/content), clarifying goals

- representation (analogies, metaphors to clarify key ideas)
- instructional selection (choosing from teaching methods and models)
- adaptation (fitting represented material to the characteristics of students)
- differentiation (they call this tailoring, but the concept is taking into consideration specific variety and needs of students in the classroom)

To transform the subject matter, effective teachers draw upon a variety of instructional approaches and strategies. Adapting and tailoring instruction requires consideration of student ability, gender, language, culture, motivations, as well as, “What student conceptions, misconceptions, expectations, motives, difficulties, or strategies might influence the ways in which they approach, interpret, understand, or misunderstand the material?” (Schulman, 1987, p. 17). Formative assessment is a powerful tool used in the process of adaptation and tailoring.

A plan for instruction results from engagement in these processes—the teacher plans and rehearses for the “performance of teaching” (Schulman, 1987, p. 17) which has not yet occurred. Instruction includes organization, management, explanation, description, work assignment, effective student-teacher interaction, and feedback. During instruction, teachers must present clear explanations and vivid descriptions, and interact effectively with students through questions and probes, answers and reactions, praise and criticism.

Schulman provides a powerful example of a teacher whose teaching style was usually interactive and flexible. However, when faced with teaching content she was unsure about, the teacher resorted to more didactic, lecture and recitation style to discourage good questions which she might not have been able to answer. Evaluation, the next process in Schulman’s model, included checking for understanding while teaching as well as more formal testing and evaluation to provide feedback and grades.

“To understand what a pupil understands will require a deep grasp of both the material to be taught and the process of learning. This understanding must be specific to particular school subjects and to individual topics within the subject” (Schulman, 1987, p. 19).

After instruction and evaluation, teachers look back on the teaching and learning that occurred through the process of reflection. Particular kinds of analytic knowledge would be combined with strategies to examine accomplishments and events that occurred during the lesson. Finally, teachers should be able to consolidate the experiential learning into a new understanding to build their professional repertoire.

To demonstrate PCK, Schulman argues that teachers should demonstrate the capacity to engage in these processes, not necessarily in sequence, and that teacher education should provide aspiring educators opportunities to develop all the processes included in the model to build deep PCK.

The term Pedagogical Content Knowledge is used widely in educational literature—Schulman was the first to articulate the notion of PCK as the interplay between pedagogy and content knowledge.

The Importance of Uncovering Student Ideas: Taking Action to Improve Instruction

Researchers at the Harvard-Smithsonian Center for Astrophysics examined the relationship between teacher knowledge and student learning, finding that, in addition to discipline-specific content knowledge, teachers needed to know common student misconceptions to realize large student learning gains (Sadler et al, 2013). The researchers administered identical pre- and post-test assessment items to teachers and

their students to find out whether teachers' knowledge of a particular science concept predicted student gains on that concept, and whether teachers' knowledge of common student misconceptions related to a particular science concept predicted student gains on that concept (Sadler et al, 2013, p. 1026).

The researchers compared pre- and post-test data from 181 middle school science teachers and 9,556 students. The multiple-choice test questions were based on the physical science content standards published by the National Research Council in 1996 (NRC, 1996). Twelve of the 20 test items had "strong" misconceptions: the item had a strong misconception if 50% or more students who chose a wrong answer preferred one particular incorrect response.

Sadler's team found that teachers on average missed only 3 out of 20 items, demonstrating 84.5% proficiency on subject-matter knowledge (SMK). However, teacher ability to identify the most common wrong answer on misconception items was weak, averaging only 5 out of the 12 items with strong misconceptions. They labeled this Knowledge of Student Misconceptions (KOSM) and then analyzed student gains in the context of teacher SMK and KOSM.

Students were aggregated into two groups: students with low reading levels and students with high math and reading scores. Students with higher reading and math scores showed much larger gains overall than students who scored low on non-science items, regardless of teacher SMK or KOSM. A clear relationship between teacher knowledge to student gains was shown: "For non-misconception items student gains are nearly double

if the teacher knows the correct answer. When items have a strong misconception, students whose teachers have KOSM are likely to gain more than do students of teachers who lack KOSM” (Sadler et al, 2013, p. 1041).

Higher non-science students seemed to benefit greatly from teachers with knowledgeable teachers. Although the gains were less significant, students with low reading and math scores exhibited no significant gains unless their teachers had the requisite SMK for these items (Sadler et al, 2013, p. 1040). “If teachers hold such misconceptions themselves or simply are unaware that their students have such ideas, their attempts at teaching important concepts may be compromised” (p. 1025).

Sadler et al (2013) emphasized that professional developers need to identify and remediate specific holes in teachers’ knowledge, since teachers must know the science and misconceptions surrounding each particular concept. For example, a teacher’s firm grasp of electrical circuits and relevant misconceptions appears to have little to do with the effective teaching of chemical reactions (p. 1041). The data also demonstrated that teachers are more effective if they know their students’ most common misconceptions. “This particular component of PCK may allow teachers to construct experiences, demonstrations, experiments, or discussions that make students commit to and then test their own ideas...It is better if a teacher also has a model of how students tend to learn a particular concept...” (Sadler et al, 2013, p. 1043).

Although Sadler et al (2013) provided evidence to suggest that teachers who are familiar with common student misconceptions are more effective, there is some evidence

to suggest that teachers have difficulty designing and implementing appropriate instructional strategies to address student needs even when they are able to uncover misconceptions or gaps in reasoning.

In *From Evidence to Action: A Seamless Process in Formative Assessment*, Heritage, Kim, Vendlinski, and Herman (2009) found that teachers are better at making inferences about student understanding than they are at deciding the next instructional steps needed based on information gathered during formative assessment. The authors analyzed the results of a generalizability study (G study) that measured teacher knowledge in mathematics. The teacher knowledge measures were part of POWERSOURCE, a formative assessment strategy for mathematics developed at the National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Teachers were asked to review student responses to assessments checking understanding of the distributive property. Using a scoring rubric, Heritage, et al., measured teachers' ability to 1) identify the key mathematical principle addressed in the response, 2) infer what the student does and does not understand, 3) decide what written feedback they would provide to help students improve, and 4) plan what to do next instructionally. One hundred and eighteen sixth grade teachers from across Los Angeles volunteered and participated in the study. The authors found that 68 out of 107 teachers received the lowest score (1) in determining the next instructional steps, even though only 13 among those 68 received a similarly low score in evaluating student understanding.

Heritage and her colleagues go on to point out, “If teachers are not clear about what the next steps to move learning forward should be, then the promise of formative assessment to improve learning will be vitiating” (Heritage, et al, 2009, p. 29). They discuss the need for teachers to have a better understanding of learning progressions, which describe the trajectory of learning along which students are expected to progress, but also acknowledge that in the United States, teachers rarely have “...adequate time to engage in deep, reflective, and ongoing discussion with each other...” (p. 30).

Although this study provided evidence to show that using assessment information to plan subsequent instruction tends to be the most difficult task for teachers in mathematics, the results revealed the significance of translating information elicited through formative assessment into the appropriate next instructional steps to improve student learning in any domain.

Teachers need to know the most common misconceptions held by their students in order to improve learning, but they also need to know what steps to take in response to the information gathered during instruction. Authors of both studies pointed out the need for further research to clarify how teachers could use formative assessment to understand student thinking and reasoning around scientific ideas, and to figure out what instructional steps to take to move students towards deeper understanding.

Living Examples of Implementation: Strategies, Approaches, and Impact on Student Learning

Citing evidence showing that formative assessment was an essential component of classroom work, Black and Wiliam (1998a, 1998b, 2004) called for a focus on the process of teaching and learning in classrooms:

Teachers will not take up ideas that sound attractive, no matter how extensive the research base, if the ideas are presented as general principles that leave the task of translating them into everyday practice entirely up to the teachers. Their classroom lives are too busy and too fragile for all but an outstanding few to undertake such work. What teachers need is a variety of living examples of implementation, as practiced by teachers with whom they can identify and from whom they can derive the confidence that they can do better. They need to see examples of what doing better means in practice” (Black & Wiliam, 1998b, p.10).

Researchers have begun to examine the effectiveness of formative assessment strategies and practices. Furtak and Ruiz-Primo conducted studies exploring teachers’ practices and student learning through formative assessment. Two studies took place in the context of a middle-school science curriculum, Foundational Approaches in Science Teaching (FAST), which was a collaboration between the Stanford Education Assessment Laboratory (SEAL) and the Curriculum Research and Development Group (CRDG) at the University of Hawai’i, Manoa (Furtak & Riuz-Primo, 2006, 2008).

In one study, the researchers compared the practice of using questions as a method of formative assessment to measures of student learning (Furtak & Ruiz-Primo, 2006). They wanted to find out how different levels of informal assessment practices related to levels of student learning. They identified whole-class, small group, or one-on-one discussions as assessment conversations, which provided teachers an opportunity to draw out and act on students’ evolving understanding. They described these assessment conversations as consisting of a four-step cycle (ESRU): 1) Teacher asks a question to elicit student thinking, 2) students respond, 3) the teacher recognizes the student’s response,

and then, 4) the teacher uses the information collected (Furtak & Ruiz-Primo, 2006, p. 205).

Furtak and Ruiz-Primo collected information from four teachers and the students in their classrooms. Teachers were asked to videotape their classrooms in every science session they taught during the 12 FAST 1 investigations, resulting in videotapes of 49 lessons across the four teachers over the four investigations. Researchers were able to transcribe 19 out of the 49 videotapes to determine which part of each transcript corresponded to discussion of results. They then focused on coding the individual speaking turns to capture the ESRU cycles in the transcripts.

Furtak and Ruiz-Primo found that the formative assessment practices of the four teachers were considerably different from each other, and that those teachers whose assessment conversations were more consistent with the ESRU cycle had students with higher performance. The one teacher whose students had the highest performance on their tests was the teacher who held the most discussions, asked the most concept-eliciting questions, and employed the greatest diversity of strategies that used information she had gained about student understanding. Of course, they acknowledged that the number of teachers was small, but that allowed researchers to transcribe assessment conversations in great detail.

In another study, Furtak and Ruiz-Primo explored how well four different types of formative assessment prompts elicited a range of middle school students' ideas about sinking and floating (Furtak & Ruiz-Primo, 2008). The authors focused on four kinds of

curriculum-embedded formative assessment prompts within the FAST Curriculum. They wanted to know which of the four types of prompts would reveal more about student thinking so that teachers could take action to improve student learning.

The researchers applied a coding framework to written responses and videotapes of discussions in the classrooms of four teachers. The prompts consisted of 1) graph, 2) predict-observe-explain (POE), 3) constructed response (CR) and 4) predict-observe (PO). In each classroom, students wrote independently in response to the prompts, and then there was a whole-class discussion where students' ideas were shared.

After careful analysis, Furtak and Ruiz-Primo concluded that diversity of students' responses in writing was not reflected in classroom discussions. They found that discussions do not elicit the full range of student conceptions that are evident in writing, and that whole-class discussions can give a teacher an upwardly skewed picture of student competence. Furtak and Ruiz-Primo (2006) also found that more open-ended written prompts such as the constructed response and predict-observe were more likely to represent all the possible student ideas in a class, while the graph and POE demonstrated student knowledge without providing insight into the substance of student thought.

The FAST curriculum developed by SEAL and the CRDG was one example of how formative assessment embedded in a curriculum could guide teachers towards better instructional practices. Sneider & Wojnowski (2013) were awarded an exploratory grant by the National Science Foundation to develop and test another model of a formative

assessment system for *Energizing Physics*, a new course developed by Aaron Osowiecki and Jesse Southwick, two physics teachers in Boston.

The development team set out to build formative assessment into the structure of the course in order to deepen students' conceptual understanding and to increase their self-confidence. They created a system incorporating two formative assessment approaches: "1) a process that occurs minute-by-minute in the classroom as students interact with each other and with their teacher and 2) a number of tools built into the curriculum, with guidelines for teachers to assess students' levels of accomplishment on each learning target" (Sneider & Wojnowski, 2013, p. 51).

Energizing Physics used the student-centered 5E learning model to embeds formative assessments such as self-monitoring through peer interaction, teacher observations, small group discussions, and teacher questioning cycles, throughout the lessons. Osowiecki and Southwick (in press) also designed quick quizzes to gather information on the specific learning targets of each lesson, which they call DYGIT (Did You Get IT?).

Osowiecki and Southwick observed pilot teachers at schools in three cities. They found that formative assessment takes time, and that it was difficult to identify specific moments of formative assessment because teachers constantly gathered information, monitored the pace, and made minor adjustments throughout the lessons (Sneider & Wojnowski, 2013, p. 21). Questions were designed to challenge common

misconceptions, and attention to student reasoning was built into the methodology of the Teacher's Guide.

Although Osowiecki and Southwick reported that their students at Boston Latin School performed much better on the quizzes and exams after they implemented the assessment tools than in previous years, the authors recognized they have only begun the process of developing a domain-specific program of formative assessment (Sneider & Wojnowski, 2013). *Energizing Physics* provides an example of a system of formative assessment that provides teachers and students with multiple opportunities to attend to the substance of student thought, modify instruction, and improve learning.

The authors and editors of these articles provide a few examples of formative assessment in action, but much work needs to be done to provide rich examples to teachers, researchers, and policy-makers. Why is the extent and nature of formative assessment in science so impoverished? Perhaps this is more a question of observation and documentation, since one could infer that teachers obviously use a variety of approaches and strategies to find out what their students are thinking. In any case, further research into the ways teachers and students use formative assessment could be extremely useful in moving the discussion from theory to practice.

Summary

Educators, researchers and policy-makers agree that formative assessment could make a big difference in improving achievement for students. Black and Wiliam (1998a, Black, et al., 2004) provided ample evidence on the effectiveness of formative assessment, which was not wholly refuted by critical reviews. Rather, critiques by Dunn

and Mulvenon (2009), and Bennett (2011) pushed the research community to develop better research methodology and frameworks. Bennett went on to argue that formative assessment must be embedded within a cognitive domain, and that measurement systems, professional development, and the larger educational system need to be factored in to truly realize the potential of formative assessment. Coffey et al. (2011) demonstrated the need to strengthen the disciplinary substance of formative assessment practice, and that formative assessment should be woven into the fabric of lessons. They supported the claim that teachers need to be prepared to be better listeners instead of just utilizing a list of strategies.

In the third section, Schulman (1987) articulated the concept of pedagogical content knowledge as both subject-matter knowledge and knowledge of pedagogical strategies. He developed a framework for the processes involved in effective teaching.

In the fourth section, two studies delved into the heart of the promise and challenge of formative assessment—since the point of finding out what students think is to adapt and modify learning activities. Sadler (2013) develops the idea of PCK further, According to Sadler, et al. (2013), students of teachers who can identify common student misconceptions learn more. While formative assessment had the potential to facilitate teacher and student awareness of misconceptions, the research by Heritage et al. (2009) contended that teachers are generally able to identify misconceptions demonstrated through formative assessment, but have difficulty making inferences about what students need in order to improve their reasoning abilities and conceptual understanding.

In the last section, studies by Furtak and Ruiz-Primo (2006) found that teachers who engaged students in discussion and other activities designed to elicit student ideas were more successful in helping their students learn. The researchers also found that written formative assessment prompts that were more open-ended provided teachers with more information about the range of student ideas in a class (Furtak & Ruiz-Primo, 2008). Sneider and Wojnowski (2013) presented a real-world example of how formative assessment was embedded into a new Physics curriculum.

The purpose of my research was to explore the ways teachers interpret and respond to the range of student ideas elicited through written formative assessment probes. The body of research in this review provided a context for the development of my research questions. While there is ample evidence that formative assessment supports higher student achievement and can improve teaching and learning, there is a great need to understand how teachers use specific formative assessment strategies and practices in the real world of the classroom. Researchers in the field do not have a clear picture of how teachers use formative assessment probes, particularly when these attempts are made by teachers who have not participated in professional development addressing how to use probes. Few studies have investigated one of the most important aspects of formative assessment: the actual ways in which teachers modify instruction in light of information gathered from formative assessment. The purpose of this research was to provide a window into the process experienced teachers go through as they put one innovative formative assessment tool, Keeley probes, to use in the classroom.

Method

Overview

The research reported in this paper draws upon qualitative research methodology of Grounded Theory to investigate the ways four elementary teachers used Keeley formative assessment probes to plan and adapt elementary science instruction to improve student learning. Specifically:

- How did teachers choose appropriate probes? What learning goals did teachers hope to address by using the probe?
- What instructional sequences did teachers envision when planning to use a probe?
- What did teachers notice when analyzing student data from a probe?
- How did teachers use the information to modify their instructional practice?

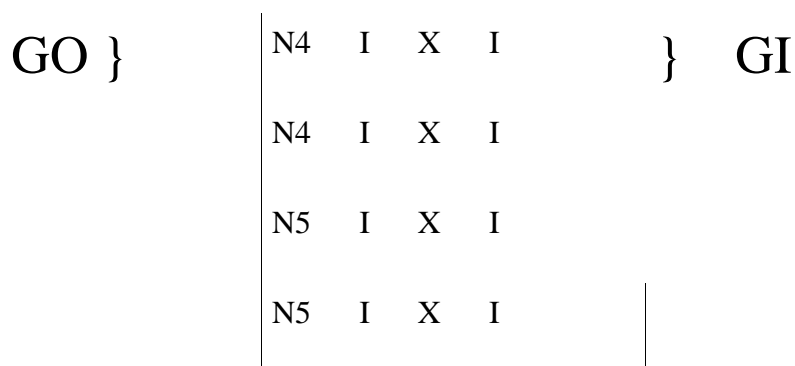
The treatment consisted of two parts: (1) introduction and orientation to the Keeley probes through a teacher-oriented professional learning community with on-going coaching for teachers, and (2) the gathering of written student responses to a Keeley probing question. Each teacher was provided with one volume of the *Uncovering Student Ideas in Science* books (Keeley, Eberle, & Dorsey, 2008; Keeley, Eberle, & Farrin, 2005; Keeley, Eberle, & Tugel, 2007; Keeley & Tugel, 2009) with ample time to review before the study began. The study commenced with a group discussion, referred to as a professional learning community (PLC) based on the Introduction to *Volume 2: Uncovering Student Ideas in Science* (Keeley, Eberle, & Tugel, 2007). The book preview and PLC discussion provided teachers the opportunity to become familiar with the

research basis for formative assessment probes and to consider suggestions for embedding probes within the curriculum.

Data was collected from two sources: 1) individual and group interviews with teachers and 2) student written responses to formative assessment probes. Notes were taken and audio captured during teacher interviews. Students' written responses were also analyzed. Because the exploratory nature of the study, I chose to use interviews as the primary instrument for gathering data.

To discover essential concepts in developing a theoretical framework, I analyzed interviews for concepts that arose and then organized concepts into tentative categories. I reviewed interview notes and audio multiple times to refine conceptual categories. I organized data into categories and developed theoretical concepts in alignment with the data.

Following is a diagram of the research design.



GO = Group Orientation/Professional Learning Community

I = Interview

X= Treatment: Keeley Probe

GI = Group Interview

N4 = Fourth Grade Teachers

N5= Fifth Grade Teachers

Participants

Data was collected from four elementary school teachers at a suburban K-5 public school outside Portland, Oregon. Two fourth grade teachers, referred to in this research by the pseudonyms Kimberly and Melissa, and two fifth grade teachers, referred to as Jason and Christopher. Participants represent a convenience sample—they were chosen based on their willingness and motivation to try a new formative assessment strategy after a series of NGSS Workshops presented by myself and a colleague from Portland State University (PSU).

Teachers volunteered to be part of the study as part of their work in partnership with the Center for Science Education (CSE) at PSU. All four teachers participated in a 2013-2014 year-long series of professional development workshops conducted by the CSE to begin implementation of the Next Generation Science Standards (NGSS) at their

school. Formative assessment was discussed during these workshops, but no specific professional development regarding the use of formative assessment probes occurred.

Although teachers' years of experience ranged from 6-18, none of the teachers had previously used Keeley probing questions. They had not participated in any professional development about how to use the probes. Teachers at each grade level had close collaborative relationships: they co-planned curriculum, analyzed student work samples together, and reflected on their practice with each other.

I also analyzed approximately 100 anonymous responses written by students. Each probe was presented as part of regular classroom instruction. These 4th and 5th graders signed assent forms, while their parents signed forms giving consent to participate. Student responses were analyzed for common misconceptions identified in the Teacher Notes.

Treatment

Keeley Probes.

This study explored how teachers plan and adapt their teaching based on student pre-conceptions revealed in probing questions developed by Page Keeley. Page Keeley, a former middle- and high-school teacher and educational leader, worked with other classroom teachers, researchers, and professional developers to write questions that could be used by teachers and students to reveal students' scientific ideas. The first volume in the series *Uncovering Student Ideas in Science: 25 Formative Assessment Probes* (Keeley, Eberle, and Farrin 2005) was published by the National Science Teachers Association (NSTA) in 2005. With the addition of the two most recent volumes

addressing astronomy (Keeley, P. & Sneider, C. 2012) and life science (Keeley, 2011), Keeley has published over 200 formative assessment probes in seven volumes, with plans for more. In this research, these probes are referred to as “Keeley probes”. Each teacher had the opportunity to choose from 100 probes total—25 in each book provided to individual teachers for a total of 100 possible choices. According to Keeley (2011),

The importance of identifying and analyzing students’ preconceptions (diagnostic assessment) and then using the data on students’ preconceptions to inform teaching and learning (formative assessment) was recognized during the 1990s in publications such as *How People Learn: Brain, Mind, Experience, and School* (Bransford, Brown, and Cocking 2000). It became clear to science teachers and educators that there was a need for student- and teacher-friendly probing questions to access students’ thinking and prior knowledge in science and to promote learning by involving students in the examination of their own and peers’ ideas. This led to the development of this popular series (Keeley, 2011, p. xi).

Keeley probes were designed with two parts: a forced-choice response where students select from a list that includes popular misconceptions, combined with a place where students are required to explain their thinking to justify why they chose a specific answer. Each probe includes teacher notes, which provide important information related to the content of the probe, national science standards, grade- level considerations, suggestions for administering the probe, commonly held misconceptions according to research, and suggestions for instruction. Keeley (2011) described how a probe works:

A probe is a specific type of question designed to reveal more than just an answer. A probe uncovers significant data about students’ thinking—for example, about their scientifically correct ideas, misconceptions, partially formed ideas, and the types of reasoning and connections they use to make sense of phenomena or concepts (Keeley, 2011, p. xii).

Each probe is accompanied by teacher notes, which include the purpose and an explanation for each probe, instructional considerations for elementary, middle, and high

school students, correlations with ideas in the National Science Education Standards, related research, and suggestions for instruction and assessment. Teachers may use Keeley probes in a variety of ways before, during, and after a lesson or activity. Each probe is designed to elicit student ideas related to a specific scientific concept.

While the probes have been field-tested, little empirical evidence exists to draw conclusions about the impact of this well-known assessment tool on teachers' instructional practices. The purpose of this study was to document the different ways that these four teachers use the Keeley probe in order to gain insights into how different ways of using the probe affect the quality of information that the teachers obtain, and how they use that data to adapt their instruction.

Professional Learning Community and Coaching.

As stated previously, each teacher was provided with one volume of the *Uncovering Student Ideas in Science* books (Keeley, Eberle, & Dorsey, 2008; Keeley, Eberle, & Farrin, 2005; Keeley, Eberle, & Tugel, 2007; Keeley & Tugel, 2009). This provided each teacher with 25 probes, creating a bank of 100 choices between the four teachers. I highlighted at least six probes in each book which would correlate with the preliminary NGSS curriculum map created by teachers during the 2013-2014 PSU Center for Science Education NGSS Workshops. Teachers received the book at the end of the 2013-2014 school year to provide ample time for review.

To familiarize teachers with formative assessment probes and establish a common vocabulary for discussion and planning, I organized a professional learning community discussion. First, I requested that each teacher read the Introduction to *Volume 2*:

Uncovering Student Ideas in Science: 25 More Formative Assessment Probes (Keeley, Eberle, & Tugel, 2007). Keeley and her colleagues wrote this introduction to describe how probes support the conceptual change model (Posner, Strike, & Gertzog, 1982), and to share suggestions for how to embed probes in instruction. By asking teachers to read and discuss the introduction with colleagues, I hoped to build teacher professional knowledge about formative assessment while also motivating them to utilize this new strategy in meaningful ways in their planning and instruction.

I also offered assistance in choosing probes and planning science units. I had established a collegial coaching relationship through the 2014 NGSS Workshops. As I spoke with teachers throughout the study, I offered suggestions to enrich the implementation of the Keeley probes within inquiry-based science units. While I challenged teachers in the study to think more deeply about the probe and instruction, the ultimate decisions about choice of probe, instruction, and follow-up were left to teachers.

In this study, teachers used three different formative assessment probes. Both 5th grade teachers chose “Is It Matter?” (Keeley, Eberle, & Farrin, 2005). Fourth grade teacher Kimberly used “Respiration” (Keeley, Eberle, & Dorsey, 2008), while her partner teacher, Melissa, collected responses to the probe “Mountaintop Fossil” (Kelley, Eberle, & Tugel, 2007). Table 1 shows a description and the purpose of each probe. Samples of each probe have been included as Appendices A, B and C.

Table 1*Description of Three Formative Assessment Probes*

<i>Probe</i>	<i>Description: Students Were Asked to...</i>	<i>Purpose: To Determine Whether Students Could...</i>
Mountaintop Fossil	<ul style="list-style-type: none"> —Evaluate four ideas and choose which they most agree with —Describe how a fossil could end up on the top of a tall mountain 	<ul style="list-style-type: none"> —recognize that mountains formed from the uplift of land, including areas that were once part of oceans
Is It Matter?	<ul style="list-style-type: none"> —check items which they believe to be matter —write a rule for the characteristics which determine whether something is matter or not 	<ul style="list-style-type: none"> —recognize forms of matter —distinguish between matter and energy, forces, and emotions —develop an operational definition of matter and identify characteristics which all matter meets
Respiration	<ul style="list-style-type: none"> —check items which they believe use the process of respiration —write an explanation of how they decided whether something respire 	<ul style="list-style-type: none"> —recognize that most living things need air to provide energy —connect respiration to breathing —recognize structures such as lungs, gills, and leaves that take in oxygen

After the group discussion helped build pedagogical knowledge between teachers, I conducted one-on-one and group interviews before probes were administered, and promptly after students responded to the probes, once teachers had a chance to review and analyze student responses. In addition, I analyzed student written responses to

compare what teachers noticed and planned with what the teacher notes for each probe suggested.

Instruments

Interview protocols were developed by the researcher, reviewed by fellow Masters' in Science Teaching (MST) candidates throughout graduate research seminars, and again by the thesis advisor. To contribute to the reliability of the data gathered from teachers, I wrote the group and individual interview questions with similar questions to provide the opportunity for ongoing reflection and the emergence of themes.

According to the Rand Corporation,

Researchers use interviews for a variety of purposes. Interviews can be used as a primary data gathering method to collect information from individuals about their own practices, beliefs, or opinions. They can be used to gather information on past or present behaviors or experiences (RAND, 2009, p. 32).

In this study, I used semi-structured interviews, sequenced from general to more specific questions as in a "Funneling Protocol" (Rand, 2009) to collect qualitative data from teachers about their formative assessment practices.

Instrument 1: Professional Learning Community Discussion Questions.

1. When you hear the words "assessment for learning" or formative assessment, what comes to mind for you?
2. What kind of interactions between students and teachers exemplify formative assessment to you?
3. At what points during the lesson cycle could teachers could use formative assessment probes?
4. For you, what are the most important goals for using assessment for learning?
5. How have you applied the conceptual change model (CCM) in your classroom?
6. Can you think of students or lessons you have taught in which students had strong ideas that were not necessarily scientifically accurate? How did you address this, or how would you now that you have had time to reflect?
7. The Introduction to *Volume 2: Uncovering Student Ideas in Science* by Page Keeley Francis Eberle and Joyce Tugel(2007) lists ten possible uses for Keeley formative

assessment probes (see page 3). Which of these uses ring true for you? Which do you feel would be most powerful to enhance student learning in science?

8. The Introduction explains the Teacher Notes that accompany the probes (page 9-10). What information is included in the Teacher Notes?
9. Is there anything else important you would like to discuss after reading the Introduction?

Instrument 2: Pre-Planning Interview Questions.

1. How do you feel about using a new strategy/tool, the Keeley Probe, in your instruction?
2. What do you usually use for assessment?
3. How do you usually know if your students have understood a concept?
4. What kind of information do you hope to find out by using this assessment probe?
5. What is the DCI/PE you have in mind for the lesson?
6. Describe the lesson sequence—where does the probe fit in?
7. How do you predict students might respond to the probe?
8. Do you plan to give feedback to your students after they have time to write their response? If so how? If not, why not?
9. Is there anything else you would like to share that was important while you were planning?

Instrument 3: Adapting Instruction Interview Questions.

1. How did you actually end up administering the probe?
2. Which responses show clear understanding of the concept?
3. Which are examples of misconceptions or varying understanding?
4. What, if anything, can you infer about your students' ability to reason effectively?
5. What else do you notice about your students' understanding from their written responses?
6. Do you plan to give feedback to your students after they have time to write their response? If so how? If not, why not?
7. After looking at what students wrote, what are your hopes for the lesson? Have your goals changed?
8. What learning activities will you provide/plan to close the gap between where students are now and where you would like them to end up?
9. How will you know your plan/adjustment worked?
10. Is there anything else you would like to share about your next instructional steps?

Instrument 4 Group Interview Questions.

1. Describe briefly how administering the assessment probe went in your classroom?
2. What kind of useful information, if any, did you find out about your students' ideas or reasoning?

3. How did you plan to adapt your curriculum/teaching in response to what you found out about student thinking with the probe?
4. Were you able to carry out the follow-up activities you planned? Why or why not? How did it go?
5. How useful were the students' written responses in helping you get an idea what your students were thinking?
6. What would you have done if you had more time?
7. How would you use the Keeley probe next time?
8. Is there anything else important that you would like to discuss?

Procedure

This study was conducted over a period of four months, with books disseminated six months prior. I collected qualitative data through individual and group interviews regarding the teacher planning cycle, analysis of student responses, and curriculum adaptations inspired by using the Keeley Probe as a formative assessment tool. The following table shows the list of activities, treatments, and instruments used.

Table 2

Research Elements and Timeline

Stage	Activity	Treatment	Instrument
Stage I	<i>Uncovering Student Ideas</i> distributed to teachers	relevant probes marked in each book, final to be chosen by teachers	
Stage II	professional learning community		Instrument 1: group PLC questions
Stage III	pre-planning interview	instructional coaching	Instrument 2: pre-planning interview questions
Stage IV	students respond to probe	“Mountaintop Fossil” “Respiration” “Is It Matter?”	

Stage	Activity	Treatment	Instrument
Stage V	post-probe interview: data analysis and instructional modifications	instructional coaching	Instrument 3: Instructional adaptation Interview Questions
Stage VI	group reflection interview		Instrument 4: Group Interview Questions

Prior to the study, I met with each teacher informally to provide a copy of one *Uncovering Student Ideas* book which included the Teacher Notes for each Keeley probe. I also explained the research questions. This afforded teachers a chance to review the Teacher Notes before the Professional Learning Community conversation and first interview, providing the opportunity for thoughtful, informed planning.

Next, teachers participated in a professional learning community discussion after reading the Introduction to the second volume of *Uncovering Student Ideas in Science* (Keeley, Eberle, & Tugel, 2007). This discussion served to provide an entry point for teachers to discuss the process of implementing formative assessment probes in their classrooms, formulate instructional goals associated with their chosen probe, and build professional knowledge about the research basis for using formative assessment probes.

Individual interviews were then conducted prior to the administration of formative assessment probes. Interviews took place one to two weeks before the lesson, in the classroom of the teacher being interviewed. If teachers did not plan to collect written responses to the probes, I requested that they do so. Teachers shared their instructional plans, learning goals, and discussed their choice of Keeley probe during the pre-planning

interview. I worked with teachers to choose a relevant probe if they had not already done so.

Teachers collected written student responses during the next stage of research. Each teacher administered one Kelley probe as part of one science unit. I interviewed teachers immediately after written responses had been collected from students, once teachers had a chance to review student responses. The purpose of the data analysis interview was to find out how teachers interpreted student ideas elicited by the probe, and to determine what, if any, instructional modifications teachers planned to make in response. By discussing and analyzing student data promptly, the hope was that teachers would have a chance to bridge any conceptual gaps in a timely manner.

During individual interviews, I also offered suggestions about curriculum and instructional strategies, and challenged teachers to think more deeply about what instruction or modifications student data might suggest. Although I had not had the chance to review the student data before the second individual interview, I provided guidance and suggestions for instruction that might further student understanding and address misconceptions, and assisted teachers in utilizing the Teacher Notes.

Finally, teachers shared their reflections and perceptions in a final group interview. The final group interview allowed teachers to reflect and share their experiences and insights conversationally. The group interview provided data about how each teacher used the information from the assessment probe to involve students in advancing their own understanding of science concepts, as well as how the experience of administering formative assessment probes affected instruction.

I analyzed the interview data by developing a coding system to identify common themes. I am offering a grounded theory addressing issues that arose as teachers used Keeley probes. Grounded theory (GTI, 2008) is a research method in which the investigator generates data by interviewing participants and then analyzes the data to determine conceptual categories. No theory is offered prior to data collection. The theoretical explanation is based on how the categories are related to each other. Russell (2000) describes the grounded theory process in six steps:

- Review interviews and read through a small sample of text
- Identify potential themes, creating analytic categories that arise
- As the categories emerge, pull together all the data from those categories and compare
- Consider how categories are linked together
- Build theoretical models using relations among categories
- Present the results of the analysis using quotes from the interviews as exemplars that illuminate the theory

To avoid bias as I developed the coding system, I took detailed notes as I listened to each interview multiple times. I removed teacher names from interview notes and re-read interviews to identify recurrent themes. I considered a theme recurrent if it arose more than four times during interviews with at least half of the participants. As themes became clear, I organized them into categories. Appendix D includes a sample of quotes used to identify and organize the themes and issues that arose during the interviews. Finally, I developed a theoretical framework to report the results of the interview process.

Results

Results are reported in four case studies followed by major themes that arose throughout the study. Each case study includes a description of the Keeley probe, curriculum and teacher process; data analysis and suggestions for follow-up instruction; and the lessons learned from the case study. In the second section, emergent themes are reported in a theoretical framework based on grounded theory.

Case Studies

Case Study 1: Melissa, “Mountaintop Fossil,” 4th Grade

Fourth-grade teacher, Melissa, chose the probe “Mountaintop Fossil” (Kelley, Eberle, & Tugel, 2007). Melissa had set a professional goal of including more formative assessment in her science instruction, and she was the most enthusiastic and motivated of the four teachers in participating in the study. In fact, it was her excitement that engaged the other teachers in full participation in the study. She was thoughtful and well-informed during interviews, embracing the opportunity to discuss how to embed formative assessment in the new geology unit she was planning.

Melissa had chosen the “Mountaintop Fossil” (Keeley, Eberle, & Tugel, 2007) probe before the first interview, and had read the Teacher Notes carefully before our discussion. She planned to use the probe to gather information before her unit on geology. She was still developing the unit at the time of the interviews. The geology unit would culminate in a class field trip to the Oregon coast to observe fossil types and rock formations directly. The geology aspect of the unit was new for Melissa, who had

previously focused on how plants and animals have adapted to particular biomes. She planned to use videos, discussion, and hands-on activities where students model the movement of rock layers. She was also searching for appropriate reading material for 9 and 10- year olds. Students in Melissa’s class would also reflect on their learning by keeping journals and portfolios.

The Keeley probe, “Mountaintop Fossil” (Kelley, Eberle, & Tugel, 2007), would provide her insight about what her students knew about mountain formation to guide her unit planning. When asked, what she hoped to find out from using the probe, Melissa replied, “I hope to find out what they understand about uplift and general plate tectonics—mountain formation. Off the top of my head, I can’t identify (the NGSS performance expectation), it’s on our unit plan, correlating as we made the outline of the unit.”

“Mountaintop Fossil” (Keeley, Eberle, & Tugel, 2007) correlated well with 4th grade performance expectations in earth and space sciences, as shown in Appendix F. The probe asked students to choose between four possible explanations for how a shell fossil ended up on top of a mountain. The best answer is Rosa: “A mountain formed in an area that was once covered by ocean.”

Student Responses

Almost half the students (11/24) in Melissa’s class chose the correct response, Rosa. However, as Melissa noticed,

In reading through responses, even kids who are close, still don’t have a deep understanding, it’s not a concept they have personal ownership over... They sort of understand how mountains are formed, but most think of lava flowing up and

out, up and out... but not much about uplift. So looking forward to that, I think there will be some big A-ha!'s!

Melissa realized that although many students chose Rosa as the correct response, students did not have much background experience with geology or plate tectonics, so their views were simplistic. A sample of student responses for each choice is included in Appendix I.

In planning the unit, Melissa had found some videos with graphics and cutaways showing how uplift happens, and giving students a visual for fissures, and the general dynamics of the earth. She had also found a few activities including one with graham crackers that show the tectonic plates colliding. Melissa seemed to welcome the opportunity to discuss the big scientific ideas raised by the probe—the formation of sedimentary rock, uplift, weathering, and erosion—throughout the interviews, but she had not yet located what she felt were the best activities and resources to help students understand these long-term constructive earth processes.

This case study illustrates the strengths and limitations of using Keeley probes to focus planning for a new science unit. Motivated and enthusiastic, Melissa sought out formative assessment strategies to boost her own knowledge of geology and deepen the conceptual knowledge of her students. Melissa felt that the grade-level articulation provided in the teacher notes combined with the specific wording in the explanation helped her hone in on what the geology standard included for her 5th graders.

She understood the importance of using multi-modal instructional strategies to address the performance expectations, however, she was left to her own devices to locate visuals, models, examples, and discussion strategies suggested in the teacher notes.

Although she felt Keeley probes would be a useful resource, she maintained that science was still “that nebulous thing for us: scrounging for materials, interpreting standards, thinking of what will match for our kids. We are still inventing the wheel as we go.”

Case Study 2: Kimberly, “Respiration,” 4th Grade

When you walk into Kimberly’s fourth grade classroom, you might notice the reading loft and bathtub reading nook, or perhaps the choice of birthday hats, crowns, and Viking helmets hanging from hooks. Kimberly’s responses were characterized by her knowledge of effective teaching strategies and her candor about her lack of science background knowledge.

Kimberly chose the probe “Respiration” (Keeley, Eberle, & Dorsey, 2008) as part of her unit about how nature provides food. This unit addressed performance expectations at both 4th and 5th grade levels, as shown in Appendix E. As part of studying where food comes from, Kimberly included what she termed “a superficial activity about photosynthesis.” Part of this garden-based learning plan included an engineering project in which students designed a plant box to increase yield in the garden. Students had also created a booklet about photosynthesis which included diagrams and an explanation of the process and structures involved in photosynthesis.

In the first interview, Kimberly was unsure whether the topic of respiration was actually essential for her 4th graders to know. I suggested she consider “Is It Food for Plants?” (Keeley, Eberle, & Tugel, 2007), which directly addressed the concept of food and how plants get energy. As an alternative, I also suggested she try “Mountaintop Fossil” (Keeley, Eberle, & Tugel, 2007) which her partner teacher chose to prepare for

the upcoming geology unit. Kimberly chose to continue with the “Respiration” (Keeley, Eberle, & Dorsey, 2008) probe.

When asked what she hoped to find out, Kimberly replied,

I want to know if they really understand the process of respiration, if they understand that it doesn't just happen in plants, but in all living things. Just at what level they understand it. We've talked and done visual notes, but I don't really know". She wanted to get at the core of what students understood, stating that photosynthesis or respiration may seem pretty simple, but understanding is actually very complex.

Kimberly planned to administer the probe after presenting the unit on photosynthesis. She planned to use the probe to find out what they knew about respiration after studying photosynthesis. She saw the probe and following discussion of respiration as a follow-up to the photosynthesis unit. During the pre-planning interview, Kimberly did not have a specific lesson or activity in mind to address the concept of respiration.

The teacher notes suggest that respiration at the elementary level is "...usually equated with breathing and focuses on familiar structures of animals and plants that take in oxygen, such as lungs, gills, and leaves" (Keeley, Eberle, & Dorsey, 2008, p. 133). For this probe, students had to put an X next to the organisms that use the process of respiration and explain how they decided whether an organism respire. The list included animals (human, fish, worm, horse, duck, honeybee), plants (grass, tomato plant, apple tree), stages of living things (frog eggs, germinating seed, butterfly larvae), and two items (bacteria, single-celled pond organisms) which elementary students might have no familiarity with.

I asked Kimberly how she would respond to the probe, and her response demonstrated confusion about the concept of respiration. “I would say... something about... eek! It has some kind of a structure inside itself that would do the exchange—that would have the cycle of... um... if they had the physical structure of being able to do the process of respiration.”

When I asked how she would define respiration in a student-friendly way, Kimberly replied, “What is respiration? Um... the exchange of... gases in or related to the atmosphere? Something that the thing needs, takes in a gas, uses what it needs, releases the end product?”

Clearly, Kimberly was having difficulty articulating what respiration is, which would make it difficult for her to engage students in scientifically accurate conversation. Kimberly was candid about her lack of scientific background knowledge, describing how difficult it was to help students understand when she is not sure what, exactly, they need to know.

When asked how she administered the probe, Kimberly reported that she followed the suggestions in the teacher notes, prefacing it with using familiar language of “Does it use air?”

Student Responses

Most of Kimberly’s students marked all the animals and plants correctly, as shown in Table 3. Three students did not mark fish, worms, grass, or tomatoes, demonstrating some confusion which might be addressed in one-on-one conversations.

Student ideas about photosynthesis and respiration surfaced as they attempted to explain how they decided which organisms respire.

Table 3

Number of Students Who Marked Organism as Respiring

Concept	Organism	# Students	% Correct
Animal	Human	24	100%
Animal	Fish	21/24	88%
Plant	Grass	20/24	83%
Plant	Apple Tree	21/24	88%
Life Cycle	Chick Inside Egg	9/24	38%
Life Cycle	Frog Eggs	10/24	42%
Life Cycle	Germinating Seed	16/24	67%

Kimberly noticed that students are “Not connecting anything that’s alive with that it’s growing and that what is growing is cells, and the cells use oxygen for that.” Since students had recently learned about photosynthesis, she also noticed that students were still confused about photosynthesis—many students thought that respiration was the opposite of photosynthesis.

Some interesting concepts also arose about the less obvious choices. Table 3 shows only 38% of students believed a chick in an egg breathes air. Students were also confused about frog eggs (42%) and a germinating seed (67%). The variety of rules students used to decide whether an organism breathes air is shown in Appendix H. Some students thought that if an organism was in an enclosed space, it would not get air, as demonstrated by the following responses:

“The ones inside of something (ex: a butterfly larvae is hibernating which requires no air.”

"I think animals in eggs don't get the air because there is a closed place.”

“I chose weter[sic] it lived outside in the open or inside an egg.”

These student ideas could serve as rich conversation starters about what is alive, and what requirements all living things have throughout the life cycle.

Responding to Student Ideas

The teacher notes for “Respiration” (Keeley, Eberle, & Dorsey, 2008, p. 132-136) suggest that the focus at the elementary level should be on respiration as breathing, and on structures such as lungs, gills, and leaves, that take in oxygen. The 4th grade NGSS performance expectation most closely related to “Respiration” (Keeley, Eberle, & Dorsey, 2008) calls for students to construct an argument that plants and animals have internal and external structures that function to support survival and growth. Appendix E provides the NGSS correlations in more detail.

After reading through student responses, Kimberly stated that she really was not sure how to go about giving feedback. She mentioned the possibility of a classroom conversation, feeling pressed for time as state testing was commencing and the end of the school year was approaching. She noted that she might include respiration in her unit on the human body next year.

One follow-up activity I suggested which might have addressed student ideas related to respiration and the relevant NGSS performance expectation was a Jigsaw activity, in which students split up in expert groups to read a short description and

analyze a diagram of how one type of organism obtains oxygen. Experts would then teach a group about the structures the organism they studied used to breathe air.

Kimberly did not end up returning to the concept of respiration. She began the geology unit with Melissa.

This case study demonstrates the difficulty even experienced elementary teachers face in teaching science concepts they do not understand. Kimberly's self-awareness about her lack of science background drove her to pay close attention to the teacher notes, yet she still had extremely limited understanding of respiration. She was not familiar with the NGSS performance expectations for her grade level, and had little idea how to address student misunderstandings. Although she discussed a variety of effective instructional strategies, she was not able to access the best practices she used in other subjects when she struggled to understand the science. She simply did not possess the scientific PCK needed to effectively teach this science concept, and the Keeley probe did not change that fact. Kimberly did reiterate that the Keeley probes provided a valuable starting place to develop curriculum based on NGSS.

Case Study 3: "Is It Matter?" Jason, 5th Grade

Although Jason assured me he was excited to try using a formative assessment probe with his students, he had not had a chance to review *Volume 4: Uncovering Student Ideas in Science: 25 New Formative Assessment Probes* (Keeley & Tugel, 2009), the book he had received the previous spring. Of the four teachers, Jason expressed the most confidence in his science content knowledge.

I met with Jason and Christopher together to discuss their probe choice and instructional plans, but, they had not decided which science unit to teach next. We reviewed the fifth grade NGSS performance expectations. Jason had just finished a unit on energy with his students, and Christopher had not presented a science unit recently.

NGSS identifies a number of performance expectations related to matter and interactions for fifth graders, as shown in Appendix G. Therefore, I steered Jason and Christopher towards preparing a unit investigating matter and interactions for their students. We found the Keeley probe, “Is It Matter?” (Keeley, Eberle, & Farrin, 2005) which would give the teachers an idea of what their students knew about matter.

At first, both teachers planned to present the probe prior to the district-provided *Full Option Science System (FOSS) Mixtures and Solutions* module (UC Regents, 2014). Because Jason felt the FOSS module took too long, I suggested he try the Great Explorations in Math and Science (GEMS) *Chemical Reactions Teachers Guide* (Barber, 1998), which includes four inquiry activities leading students to explore chemical changes by mixing chemicals that bubble, change color, get hot, and produce gas, heat, and odor.

Jason adapted the unit, written for grades 6-8, for his 5th graders, to focus on observation, experimentation, and inference. Jason was not sure how well “Is It Matter?” (Keeley, Eberle, & Farrin, 2005) matched up to the lessons he planned to do. When asked what ideas he wanted to elicit with the probe, he answered, way to see if the tool matches up with the lessons we will do. I’m not sure that’ll work in

When asked if he knew the performance expectations met by the unit, he pondered, e if the tool matches up with thethem, technology is very handy.”

“Is It Matter?” (Keeley, Eberle, & Farrin 2005) asked students to put an X next to things they considered to be matter, and then explain their thinking by describing a rule they used to decide whether something is or is not matter. The teacher notes (Keeley, Eberle, & Farrin 2005, p. 80-81) explain how this probe may be used to find out if students can recognize different forms of matter; whether they can distinguish between matter and energy, force, or emotions; and what characteristics would be used to decide if something is matter.

Student Responses

To analyze student responses, I grouped items into the following categories: forms of matter (solids, liquids, gases); forms of energy; and forces. Solids include rocks, salt, and planets. Liquids include milk, water, and dissolved sugar. Gases include oxygen and air. Heat, light, sound waves, and electricity comprise the forms of energy group. Forces include gravity and magnetic force. NGSS 5th grade performance expectations do not emphasize understanding matter at the atomic or cellular level.

Table 4

Number of Jason’s Students Who Marked Substance as Matter

Concept	Substance	# Students	Percent Correct
Solid	Rock	12/25	48%
Solid	Planet Mars	17/25	68%
Liquid	Water	12/25	48%
Liquid	dissolved sugar	6/25	25%
Gas	Air	23/25	92%
Energy	Heat	8/25	32%
Force	Gravity	9/25	36%

As Table 4 shows, most students in Jason's class recognized oxygen (21/25) and air (23/25) as matter, yet less than half marked rocks (12/25 or 48%) or salt (11/25 or 44%) as matter. This showed a common misconception that solids were not matter. Only about 25% (6/25) guessed that dissolved sugar was matter. Jason predicted correctly that most students did not believe planets and stars were matter (17/25 or 68%).

Most students thought heat was matter (17/25, 68%), while more than half thought other forms of energy: light (14/25, 56%), sound waves (13/25, 52%), and electricity (14/25, 56%) were matter. These choices show that students generally could not distinguish between matter and energy. Only 9/25 (36%) knew that gravity was not matter, but most (20/25, 80%) did not mark gravitational force as matter.

Jason did not analyze student data from each of these categories, nor did he identify a major misconception--most of his students did not mark common substances like rocks or dissolved sugar as matter.

Most likely because the class had just studied energy, Jason did notice that more than half of his students thought heat was matter. He decided to address this misconception. He conducted numerous whole-class conversations, and brought in a magazine article and video segment to develop student understanding about heat as energy.

The second part of the probe asks students to explain their thinking by describing a rule they used to decide whether something is or is not matter. According to the teacher notes (Keeley, Eberle, & Farrin 2005, p. 81), substances must meet the following criteria to be considered matter: be made up of particles, have weight and mass, take up space,

and exist in the form of solid, liquid, or gas (elementary students are not expected to discuss plasma).

Did students use some or all of these criteria to decide whether something is not matter? Samples of student responses are included in Appendix J. Jason pointed to a few responses such as “matter has mass”, “matter is made of molecules”, and “something that fills something up” to show students who had some more accurate ideas about matter.

Jason reported how the concept presented in the Keeley probe prompted discussions in his classroom:

I didn't think it (“Is It Matter?”) matched well with the activity, but it uncovered those misconceptions. It was very clear who had some misunderstanding. It changed the way I taught the lesson. It opened my eyes to how the lesson, which was more designed around chemical interactions, could focus on matter itself and what that is. (The probe) elicited great discussions, debate on what matter was and wasn't, and a great discussion about energy.

This case study provides an example of how effectively Keeley probes may nurture rich classroom conversation, and how the concepts presented in the probe help frame instruction around the big ideas in science. Jason's experience with the probe also reinforces the vast difference in PCK between elementary teachers, and the need to take this into account in designing individualized professional development and coaching opportunities for teachers. Jason felt confident about his subject-area knowledge, effectively addressed student misconceptions about heat being a form of matter, yet still struggled to find the time and best strategies to provide multiple opportunities for students to develop conceptual knowledge.

Case Study 4: “Is It Matter?” Christopher, 5th Grade

Initially, Christopher listed a few probes which seemed interesting, without thinking about how they would connect to the curriculum. As mentioned before, after consulting the NGSS for fifth grade with Christopher and Jason, we chose the probe, “Is It Matter” (Keeley, Eberle, & Farrin 2005) to start a unit on matter.

Christopher decided to present the district-adopted Full Option Science System (FOSS) Module *Mixtures and Solutions* (UC Regents, 2014). The module included four investigations that introduce students to fundamental ideas in chemistry. In the module, students had the opportunity to:

- Make and separate mixtures, using screens, filters, and evaporation.
- Measure solids and liquids to compare the mass of a mixture to the mass of its parts.
- Use a balance to determine relative concentration. Layer solutions to determine relative density (concentration).
- Plan and conduct saturation investigations. Compare the solubility of substances in water.
- Identify an unknown substance based on the properties of solubility and crystal form.
- Observe and compare reactants and products of several chemical reactions.

Christopher reported that he had presented the FOSS unit before, but had never made it through all four investigations.

We discussed a few possibilities for integrating the probe into the curriculum. I suggested an activity called coming to consensus, in which students would develop an operational definition of matter in cooperation with classmates. I also suggested providing students an opportunity to re-examine their ideas after observing properties of matter directly and investigating mixtures and solutions through the inquiries in the FOSS module.

Christopher decided to administer the probe before presenting the unit, and then administer it again once students had completed a few of the investigations. He felt the probe might be a way to anchor the teaching he was doing, to prevent getting side-tracked by the plethora of concepts and conversations which might be raised throughout the unit.

Student Responses

To explore student ideas regarding states of matter, I used the same categories to analyze responses from students of Jason and Christopher. Table 5 shows the number and percentage of students who identified each substance as matter. Most students in Christopher's class recognized that water (21/23), air (20/23), and planets (19/23) were matter, but were fairly split on whether solids (rocks 12/23, salt 11/23) were matter. A few more than half (14/23, 60%) considered dissolved sugar matter. Most students knew that heat (16/23) and sound waves (20/23) were not matter, yet only 4/23 knew that light was not matter. Most students also knew that the forces of gravity (18/23) and magnetic force (19/23) are not matter.

Table 5

Number of Christopher' Students Who Marked Substance as Matter.

Concept	Substance	# Students	% Correct
Solid	Rock	12/23	52%
Solid	Planet Mars	19/23	83%
Liquid	Water	21/23	91%
Liquid	dissolved sugar	14/23	60%
Gas	Air	20/23	87%
Energy	Heat	7/23	30%

Concept	Substance	# Students	% Correct
Force	gravity	5/23	22%

In Christopher’s class, eight students included “something you can feel” or “if you can hold it” in their rule. Appendix K shows a sample of the rules students described, from naive to more scientifically accurate. When asked what most stood out for him after reading students responses, Christopher wondered about the ability of his students to generalize. “This question, ‘Describe the rule’, that was a different way of thinking for them. A couple kids that are more scientific thinkers, they tried, but most of them had a hard time generalizing, they don’t understand how to generalize, or find a criteria...” stated Christopher.

“What a lot of kids wrote, well, there’s a sense of what’s matter, but not necessarily a scientific sense.” Christopher picked responses like, “Something you can feel,” and, “If you can hold it or part of it in your hand without changing shape, then it is matter” as the most accurate responses.

Responding to Student Ideas

Christopher did not plan additional activities based on the ideas revealed in the probes, but he did plan a follow-up conversation in which students looked back at their initial responses as he conducted a whole-class discussion about which items were matter. Students used colored pencil to mark any responses they had incorrectly marked. In this way, he hoped to encourage student self-awareness and engage meta-cognition.

Christopher believed he could address some of the less developed thinking about matter by working through the FOSS curriculum. He noted that the next investigations

focused on saturation and chemical changes, giving him an opportunity to develop the idea that just because you cannot see it doesn't mean it is not matter. He planned to provide ample opportunities for students to do a lot of writing to "check into their mind, to make them more accountable." He stated that the probe helped anchor student thinking and made students more aware of what they know so that they might be able to focus on what they still need to know.

The teacher notes (Keeley, Eberle, & Farrin, 2005, p. 82-83) for this probe suggest providing students with opportunities to observe the three states of matter, differentiate between physical and chemical changes, and classify elements, compounds, and mixtures. The FOSS module includes these kind of inquiry activities. To address the specific confusion students demonstrated about solids like rocks and salt, I suggested a brief observation-based activity in which students identify properties of common solids like salt, sugar, and rocks that he could bring in, followed by a Think, Pair, Share, such as, "Do rocks take up space? Do they have weight?" Students could have incorporated this experience into the preliminary operational definitions created in groups.

Christopher did not end up completing the FOSS module. Instead, he started a new activity called BizWorld, which he was more enthusiastic about teaching. He repeatedly expressed his appreciation for the way the Keeley probe helped him anchor teaching around the big idea of what matter is. "I liked it, too, because it helps students, and it helps me...it prompted good discussions and made them think about, 'What is matter?' It wouldn't necessarily occur to them to think about, 'Why is fire matter?'"

The final case study points again to the challenge teachers face as they struggle to plan and teach science units on subject matter they are unfamiliar with, particularly in light of the limited amount of time allocated for planning combined with the wide array of demands faced by elementary teachers. Christopher appreciated how the Keeley probe got right to the heart of student ideas and misconceptions, and felt the probe was extremely effective at providing an anchor to important scientific ideas. With limited time, Christopher acknowledged that he chose to use the FOSS module because he liked not having to create everything from scratch. As a busy teacher and father of young children, he expressed the desire for a more efficient way to get to the core ideas he wanted students to understand. Like the other teachers in the study, Christopher often felt that NGSS performance expectations were dense and difficult to unpack. Ultimately, although he felt the Keeley probe helped teachers and students be more aware of important concepts, he did not modify instruction to address misconceptions, and abandoned the matter unit before completion.

CONCEPTUAL FRAMEWORK: Recurring Themes

I developed the theoretical framework based on recurring themes that arose as teachers reflected on the process of eliciting student ideas with Keeley probes. Appendix D highlights the conceptual codes used to organize interview data.

Theme 1. Learning Goals and Teacher Pedagogical Content Knowledge

- Teacher Domain Knowledge
- Standards Alignment
- What is Grade-Level Appropriate?

Theme 2. Instructional Decision-Making

- Anchoring Instruction with Big Ideas
- Knowing How to Address Misconceptions
- Leading Classroom Conversation and Providing Feedback
- Finding Time for Multiple Forms of Instruction

Theme 3. Data Analysis and Interpretation

- Analyzing Student Responses
- Knowing How to Address Misconceptions

Theme 4. Student and Teacher Engagement

- Promoting Student Self-Awareness
- Teacher Attitudes About Formative Assessment
- Student Engagement

Theme 1. Learning Goals and Teacher Pedagogical Content Knowledge

A. Teacher Disciplinary Knowledge

Quite often, teachers, especially elementary teachers, do not have the disciplinary knowledge needed to clarify and deepen student understanding (Coffey, et al, 2011; Sadler, et al, 2013; Black, et al, 2004). In fact, all the teachers in this study expressed doubt in their scientific understandings. They also felt the Keeley probes were useful to develop more understanding of the specific concept they needed to teach. As Melissa stated, “I like how these probes give a place to start...start searching, review the basic concepts. If I personally have misconceptions I can do some research myself to present accurate information.”

Jason worked to understand the specific content knowledge involved in the *Chemical Reactions Teacher Guide* (Barber, 1998). “So, really, what exactly is going on

in this reaction? What is the gas that is being produced? That's where the heat is coming from, right?"

Sadler, et al (2013) claimed that it is important to examine teacher knowledge surrounding particular concepts, because "...Student performance at an item level is associated with teacher knowledge of a particular concept" (p. 1041). According to Black and his colleagues (2004), however:

A high level of qualification in a subject is less important than a thorough understanding of its fundamental principles, an understanding of the kinds of difficulties students might have, and the creativity to be able to think up questions that stimulate productive thinking (Black et al, 2004, p. 17).

Thus, in this study, Keeley probes provided much-needed support to teachers in clarifying their own misconceptions and preparing them to communicate more clearly with students. However, teachers still struggled to understand scientific concepts, as Kimberly articulated, "I don't have a science background; it's hard to teach to that conceptual level. I feel like, you are walking around something, but not getting to the real thing."

B. Standards Alignment

Christopher talks about the challenge of unpacking science standards:

Also, the standards, well, content-related ideas that they say, 'teach this', but it's dense. So if I'm not reading closely, or if I don't understand, I will miss this, and I think I have in the past. So it's there, it's helping me see it and make it more clear to students.

Kimberly concurred, "I can go to the standards, but that doesn't always help me. This (the Keeley probe) helps unpack it for me, what are they really asking? What is the underlying concept?"

C. What is Grade Level Appropriate?

Melissa expressed the challenge of understanding what students at each particular grade level need to know:

I was looking at some of the other ones (formative assessment probes), as a 45-year-old, I am having to think, I'm not quite sure! So to be able to read through those grade band differentiations, my ten year olds are supposed to be able to know this but I haven't been very clear about it.

Kimberly explained that the teacher notes helped her decide how in-depth to go, since students wanted to know more and she was not sure how much they really needed to know about respiration in 4th grade.

Interviews revealed substantial information about how teacher PCK played into the use of the Keeley probes, and how well teachers were able to connect the big scientific ideas raised by the probe to standards-based learning goals.

Theme 2. Instructional Decision-Making

A. Anchoring Instruction with Big Ideas

Teachers in this study found that they were more able to focus on the big ideas throughout their instruction when they started the unit with a Keeley probe. According to Jason:

It [Is it Matter?] was very useful in thinking about what to do with Chemical Reactions. Well, it's not really what the lesson was about, but it really helped me to frame what was happening, helped solidify what was happening in the reactions and why we were observing.

Christopher continued.

It was the anchor, the conception that you want to change and help articulate. It was kind of nice to start with that. Yes, we are dissolving sugar and salt, yeah, um...ok. But there is something more we need to be aware of.

Melissa had a similar viewpoint.

Like, I understand at a superficial level, but, because the wording and concept is more specific in the probe, I can be more clear about what is actually the big idea—mountains formed over time because of uplift! As opposed to um, sort of, this is kind of what happened.

So, Keeley probes helped teachers in this study focus on the big scientific concepts. Christopher articulated how the probe anchored his lesson:

The probe was efficient, got right to the misconceptions or no conceptions. Describe the rule—that honed in on the generalization—can you make a generalization? The topics are good, the possibilities are tricky—it invited good conversations. It helped me as I go through the curriculum to work towards a more enduring understanding. Essential questions—it anchored my thinking as a teacher and helped me as a teacher focus.

B. Knowing How to Address Misconceptions

Kimberly articulated how many teachers feel about addressing student misconceptions when she said, “Really, I am just not exactly sure how to go about it”.

Jason discussed the process he went through to address misconceptions about heat and matter.

Half and half, we are still going back and forth on heat. Heat—well, some said, it must be matter because you can feel it, they really struggled with that. It’s been like 3 to 4 days, so today we watched *Magic Schoolbus* at lunch, had a discussion about it. No one could articulate why it’s not matter until we related it back to sound and sound waves. Finally, one kid said, ‘Sound is energy!’ Then we talked about electricity from last year, then insulation, then finally, they thought, heat is energy. Now I’m pretty sure they all understand... heat anyway.

When students had to decide whether heat was matter a second time, only three students (3/25 or 12%) still marked heat as matter (compared to 17/25 or 68% the first time). Jason had found a way to address this misconception, but research shows many teachers struggle to make inferences about what students need in order to improve their

reasoning abilities and conceptual understanding (Heritage, et al, 2009; Sadler et al, 2014).

Christopher addressed this challenge as well. “I was hoping for more of, you know, ‘This is what you can do to teach them, more of, if they have this misconception, this is a suggestion.’”

Kimberly concurred.

Do these people have like a connected, like, where to go after this? Because it makes you think about what you are not getting across, but I don’t always know the next really good step, it just needs to be straight-forward and clear, and discovering it versus, you know, ‘I’m just going to tell you how this works.’ Because it doesn’t work that way. Like, where could I go to get that idea for an activity and discovery instead of just telling them something.

Therefore, this research revealed that teachers do not always feel they have access to the follow-up activities, visuals, and well-thought-out discussion questions they need to address student misconceptions.

C. Feedback: Leading Classroom Conversations

In this study, teachers reported using classroom conversation as the primary form of feedback to their students after eliciting ideas with the probes. Kimberly shared her plan to use conversation as feedback: “I don’t know if I’ll give specific feedback, I’ll probably generalize and then open up in a conversation.”

“Listening is important, too. What is their perspective? Looking for what they understand...” continued Christopher.

Furtak and Ruiz-Primo (2006) adopted the term *assessment conversation* to refer to the daily instructional dialogues that embed assessment into an activity already occurring in the classroom (p. 207). They characterized these assessment conversations

by a four-stage ESRU cycle. The teacher elicits student thinking by asking a question, the student provides a response, the teacher recognizes the students' response, and then uses that information to support further learning (Furtak & Ruiz-Primo, 2006).

Teachers in this study reported using this type of conversation to improve student learning. Although this study did not include classroom observations, teachers felt they had better classroom conversations using the information they gathered from the Keeley probe. Jason went on to say,

You could really use any of the probes, you would still get the rich discussion with whichever you chose. The combination of the activity and the probe produced the conversation, but that became the basis for understanding what matter is. That would not have happened without the probe, you use that as the frame.

Keeley calls these “Juicy Questions”—questions that require students to think deeply and extract knowledge that will help them answer a rich, novel question (Keeley, 2008). Teachers in this study certainly viewed Keeley probes in this light.

Kimberly, however, noted the need to move beyond conversation:

The probe helped me shore some things up. We've got a unit we have been developing, but we haven't done much with the geology piece of it. It helped me hone in on what is geology at this grade level, how is this piece connected to what we are doing, and am I providing enough discussion, examples, visuals, models and activities, as opposed to glossing over it with one conversation.

D. Finding Time for Multiple Forms of Instruction

Teachers agreed that finding time to plan and present a variety of learning opportunities addressing performance expectations and student ideas revealed by using the Keeley probe was a big challenge. Teachers need time to review and understand the scientific concepts, choose the best activities and materials, and become familiar with

possible misconceptions. Students and teachers also need time in class to engage, explore, explain, elaborate, and evaluate. Limited planning and instructional time do not always allow teachers to respond to student ideas effectively.

According to Jason, “The probe also made it clear, still clear, how much experience in multiple ways students need to really build concepts. That is always a surprise.”

Bennett (2011) also identified time as a substantial challenge in his discussion of professional development.

Even if we can find a practical way to help teachers build pedagogical skill, deep domain understanding, and a sense of measurement fundamentals, teachers need significant time. They need time to put that knowledge, skill, and understanding to practice, for example, to learn to use or adapt purposefully constructed, domain-based, formative-assessment materials” (p. 19).

It seems learners of all ages need significant time to develop authentic conceptual understanding.

Christopher summed up the challenge he saw with this. “You were talking about time. The activity, the discussion, drawing and writing and reflecting, that’s a couple days. The writing—they need time to do that. I don’t always give enough time for that reflection to cement the learning.”

Theme 3. Data Analysis and Interpretation

Analyzing Student Responses

Teachers in this study did not analyze student responses for specific concepts based on information from the teacher notes. For example, Jason generalized about student ideas, but did not identify whether students understood the specific criteria, such

as solid, liquid, or gas, used to determine whether a substance is matter. Jason had this to say about student responses:

My student responses ranged from those who wrote a rule that is the absolute opposite of what matter is to pretty close. But, those that were close to the definition with their rule, did not choose the right items as matter. Their x's did not really match up. There weren't that many of them that had a clear conception.

Melissa also made general observations about student responses without referring to specific conceptual understanding. "They just don't have much background experience with geology or plate tectonics, so their views are very simplistic. The hands-on piece and the visuals will be helpful, as well as the big picture of how plate tectonics work."

Knowing How to Address Misconceptions

Although one teacher held a classroom conversation to address specific misconceptions students expressed about heat and matter, in general, teachers in this study did not plan specific learning activities in response to information from Keeley probes. Christopher felt he could address naïve scientific ideas by working through the box curriculum. Kimberly candidly expressed her uncertainty at how to go about helping her students better understand respiration. Melissa planned the geology unit with multi-modal instruction in mind, but was stymied when asked what learning activities she might use to close the gap between what they had written on the probe and what she wanted them to understand about mountain formation.

Teachers in this study did not seem aware of the importance of providing feedback or following up with specific misconceptions expressed by students. My impression during preliminary interviews suggested they were interested in getting a general idea about the range of student understanding, but had not considered modifying

their original unit plan to include specific follow-up discussion and activities based on the unique responses to the Keeley probe.

Theme 4. Student and Teacher Engagement

A. Promoting Student Self-Awareness

“Jason and I have enjoyed thinking about using formative assessment not just as a way to anchor science instruction, but as a way to anchor kids thinking and them being aware of what they know, helping kids focus in on what they need to know,” reported Christopher.

Jason put it this way, “Maybe they will be paying attention to what they are learning (about matter), instead of just, ‘Woo-hoo! It’s expanding! It’s getting hot!’”.

Student self-assessment is fundamental to the power of formative assessment to improve learning. Students learn more when they are active participants, and when they are able to assess their own level of understanding (Sadler, et al, 2014; Black & Wiliam, 1998a). Black and Wiliam (1998b) identify the development of student self- and peer-assessment skills as a key component of formative assessment.

Teachers in this study reported that using Keeley probes helped students develop the capacity to work at a metacognitive level. Melissa prompted students to re-visit their first response to the Mountaintop Fossil probe later in the unit, and then figure out what they had learned. “That’s kind of a big ‘A-ha!’ for them, to see, oh, I used to think that! Oh my gosh!”

B. Positive Teacher Attitudes Towards Formative Assessment

All the teachers in this study were enthusiastic and thoughtful in using Keeley probes to tackle the real work of implementing meaningful formative assessment. They all recognized the potential power of formative assessment to improve student learning. While discussing his plans for the probe, Christopher stated, “Truthfully, I am excited about the probe, getting them to demonstrate their thinking in a different form, so it will be interesting to see what they have to say. I’m curious.”

Jason concurred. “I look forward to seeing if it tunes me in, and even more if it tunes the kids into the lesson—into what they are supposed to be learning and thinking about.”

Melissa concluded, “I am not good at using formative assessment yet, I am hoping to use more around science and historical ideas as well. I look forward to using more.”

C. Engaging Student Interest

Keeley probes captured student interest in new science topics, according to teachers in this study. “The students wanted to get it right, but it wasn’t scored. They wanted to figure it out, what is matter? So it was really kind of what you want education to be,” reported Christopher.

The probe also sparked student interest in Melissa’s class. “Students were interested, motivated. It never occurred to them that they could find a shell fossil on top of a mountain, they had never even thought that was possible...I gave it a full six weeks ahead, so it’s been a while. Kids have asked when we’ll talk about the fossil on the mountain,” explained Melissa.

Jason expressed surprise at how engaged his students were.

The probe lent itself to have a discussion right away with them, it seemed to pique their curiosity. I was sort of surprised by that, that they were very engaged and interested. They were all engaged in the discussion. All of them!

Therefore, the Keeley probes used in this study effectively sparked student interest and set the stage for further scientific inquiry.

Summary of Key Findings

1. Teachers in this study reported that Keeley probes helped anchor teaching and focus on essential questions.
2. Teachers in this study primarily chose to use Keeley probes before beginning a new science unit. They had difficulty choosing probes aligned with grade-level standards.
3. Teachers used the probe to guide whole-class discussion throughout the pre-planned unit, or used the suggestions from the Teacher Notes to guide unit planning. For the most part, teachers in this study did not plan additional follow-up activities to address specific misconceptions revealed in the probes.
4. Teachers in this study had questions about the disciplinary core ideas in the units they were teaching. Most expressed some lack of confidence in their pedagogical content knowledge. Teachers reported that the Keeley probes and the accompanying Teacher Notes helped them shore up content area knowledge and focus on big ideas.
5. Teachers reported a high level of student engagement with the scientific concepts presented in the Keeley probes.
6. Teachers in this study expressed the need for support in finding efficient, multi-modal follow-up activities that would address specific student misconceptions. They

- reported feeling unsure about how to address inaccurate student ideas uncovered by using the Keeley probes.
7. Teachers in this study held positive attitudes toward formative assessment, hoping to incorporate more formative assessment into their science teaching.
 8. Every teacher in the study mentioned time as a limiting factor in their ability to fully implement modifications or multi-modal instructional activities needed to deepen student conceptual understanding.

Teachers in this study used formative assessment probes primarily to anchor their instruction to a big scientific idea. They used probes to gather information about student ideas before presenting a unit. The primary form of feedback they gave students was in whole-class conversations throughout the units. Some teachers needed to review the pedagogical content knowledge or felt unsure about the science concepts presented in the units. The Keeley probes seemed to help teachers identify and understand which key concepts to focus on at their particular grade level, but teachers still struggled to address appropriate grade-level performance expectations without getting off track. While the use of the Keeley probes provided teachers with useful information about student thinking, this research demonstrates that teachers need time and training to become masterful in the regular use of formative assessment techniques and tools.

Discussion

Limitations

Broad generalizations about the way teachers use formative assessment probes would most likely be inappropriate due to the nature of the data and small number of teachers participating in the study. Participating teachers volunteered to join this project, so results may not be generalizable to other elementary school teachers. Their insights may also have been affected by their participation in the NGSS workshops conducted by the Center for Science Education at Portland State University, or by their previous interactions with the researcher as a professional developer and instructional coach.

Another concern is the nature of how data was gathered in this study. Interviews were held after school, at a time when teachers were often exhausted and working to prepare for the next day while looking forward to ending the work day. Because this study relies on teachers to self-report, it is possible that essential elements of planning, instruction, or reflection were left out by teachers simply because they did not have substantial time to decompress and reflect.

However, since the purpose of the research was to shed light on putting formative assessment into practice in the real world of the classroom, I offer herein a few theories worthy of consideration and future research.

Answers to Research Questions

How did teachers choose appropriate probes? What learning goals did teachers hope to address by using the probe?

Teachers in this study had difficulty choosing probes aligned with NGSS, or any science standards. Only one teacher had chosen a probe with a specific NGSS performance expectation in mind. The other teachers expressed uncertainty about which probe to choose and whether it was aligned with grade-level expectations.

Three out of four teachers in this study did not identify student outcomes or performance expectations when discussing their unit plan. When asked what learning goals guided planning, most listed concepts, like “living/non-living things have characteristics” or “learn about solutions and mixtures”. One teacher was able to articulate what students should be able to do: students would notice what happens when you mix chemicals in the baggie during the first *Chemical Reactions* (Barber, 1998) inquiry, and then design a procedure to decide which chemical produces the heat in the reaction,

When asked a follow-up question about NGSS, teachers in this study were able to find the relevant standards, but still did not seem to understand how the performance expectation, the format by which students would demonstrate understanding, should be driving planning and instruction. I worked with each teacher to analyze and break-down the relevant performance expectations, but many questions remained in teacher’s minds.

Sadler (1998) discussed how criteria teachers bring to an assessment task exist in an unarticulated form, which makes them difficult to share with learners, or in a standards-referenced form (p. 3). In this study, criteria were mostly unarticulated by teachers. In fact, I would tentatively conclude that teachers actually paid more attention to NGSS because of participating in the study—I pushed them to locate and articulate the

performance expectation during the planning interview. Once they had examined the disciplinary core idea and performance expectation linked to the Keeley probe, every teacher expressed difficulty unpacking the standard.

Therefore, evidence from this study suggests that unless elementary teachers have access to instructional support in science, they are likely to plan science units or activities without aligning to NGSS in more than a superficial manner. Furthermore, data suggests that, without guidance or collaborative planning time, these teachers use Keeley probes randomly, haphazardly linked to unarticulated learning goals. Of course, students still had meaningful opportunities to learn real science in these classrooms, but without connecting formative assessment to clearly articulated learning goals, students most likely did not learn as much as they could have. Although teachers reported that using Keeley probes helped focus instruction on big ideas, the results described above also point to significant barriers in implementing formative assessment effectively.

What instructional sequences did teachers envision when planning to use a probe?

I wanted to know how the Keeley probe would fit into the way teachers organized learning activities. The order and organization of learning activities affects how well students learn and remember. I also wanted to know whether teachers used the suggestions provided in the teacher notes.

Three out of four teachers used probes to gather information about student ideas prior to instruction, while one teacher used the probe as a follow-up after instruction on a related topic. In every case, teachers did not integrate formative assessment as an ongoing aspect of teaching and learning, but, rather, presented the Keeley probe as a discrete activity, whether once or twice.

Both fifth grade teachers chose a “box” curriculum and followed it closely. One fourth grade teacher collaborated to create a geology unit based on NGSS, using suggestions from the teacher notes and information from “Mountaintop Fossil” (Keeley, Eberle, & Tugel, 2007) to inform her lesson planning. She set a goal of providing ample visuals, models, and opportunities for students to build understanding through direct experience with landforms. She also mentioned one-on-one conversations that might be built into the field study as parents reviewed journals with students.

During the pre-planning interview, three out of four asked for suggestions about how to use the probe during instruction. I pointed out suggestions in the teacher notes and offered additional suggestions, including the possibility of giving students a chance to reexamine their initial responses on the probe in some way, and directed each teacher to the teacher notes for more ideas. After discussion, three out of four chose to present the probe a second time, giving students the opportunity to self-assess and document any changes in understanding. Although none of the teachers prepared targeted questioning sequences, they did report conducting “on-the-fly” class discussions to clarify some of the misconceptions exhibited on the probes.

These results suggest that, although teachers were experienced and may have possessed professional knowledge about pedagogy, when faced with implementing this new formative assessment strategy during science instruction, they needed prompting and support to plan interactive, discussion-based learning activities. None of the teachers used flexible grouping, jigsaw, think-pair-share, or other well-known best practices for encouraging student scientific discourse.

What did teachers notice when analyzing student data from a probe?

According to Coffey, et al (2011), “The core of formative assessment lies not in what teachers do but in what they see. The point is teachers’ awareness and understanding of the students’ understandings and progress: that’s what the strategies are for” (p. 1128).

In general, teachers did not analyze responses to the formative assessment probes using the explicit purpose and explanation included in the teacher notes with each probe. Observations made were very general. One teacher effectively identified the confusion students had about whether heat was matter, and conducted follow-up instruction and discussions to help students explore this concept. In general, teachers in this study were able to articulate which students demonstrated naive understandings and which had more sophisticated conceptual knowledge, but they did not infer what specific concepts were confusing or what actions they might take to move students towards deeper understanding.

How did teachers use the information to modify their instructional practice?

All the teachers in this study reported using the ideas presented in the probe to engage students and to focus whole-class discussions throughout their pre-planned science unit. As Jason explained, “It was very useful in thinking about what to do with the *Chemical Reactions* unit. Well, it’s not really what the lesson was about, but it really helped me to frame what was happening, helped solidify what was happening.”

All four teachers reported using “on-the-fly” responsiveness in assessment conversations (Furtak & Ruiz-Primo, 2006) to further student scientific understanding.

Since classroom observation was not a part of this study, data regarding the quality and content of these informal conversations is not available.

While all the teachers in the study used data gathered in the probes during pre-planning instruction, most did not adapt instruction. That is, they continued with the plan they had made, even when given suggestions for possible follow-up activities or resources. When I suggested specific learning activities which might address some of the misconceptions, three out of four teachers stated that they simply did not have enough time to go back and present follow-up discussion or activities. The fourth teacher had not yet presented the geology unit at the time of the final interview.

This is not to say that quality instruction did not occur—teachers presented meaningful, engaging learning opportunities for students, and many students demonstrated significant scientific understanding in the second administration of the Keeley probe. What the results suggested is that teachers did not make instructional decisions based on information revealed in the probes, undermining the power of these formative assessment tools to improve student learning.

None of the teachers in this study provided individual feedback to students. It is interesting to note that two out of four teachers mentioned scoring when asked about feedback. Christopher, for example, responded,

I don't think I would score this, I don't think that would be fair, because it's information gathering... but it would also undercut what we're trying to do which is get honest ideas. I'll administer it, I'll collect it, we might talk about different answers as a class, but then we can go back to some of the misconceptions as they go through the unit and I can keep coming back to some of the misconceptions they might have had.

Every teacher mentioned whole-class discussions during and after scientific investigations. “I think of conversations, asking a specific question and then looking for what they understand” reported Jason. Both Christopher and Jason presented the probe, “Is It Matter?” (Keeley, Eberle, & Tugel, 2005), a second time to give students a chance to think about their own ideas and identify how their thinking might have changed to developing student self-assessment skills.

Quite possibly, teachers were simply not aware of the importance of analyzing and addressing specific ideas related to core scientific ideas and practices. Since this was the first time they used Keeley probes, they did not seem to realize the importance of timely feedback or instructional changes based on student ideas.

Relevance to Prior Research

The research questions provided the starting place to gather data on how teachers use Keeley probes to elicit student ideas. The more significant contribution made by this research is to shed light on how formative assessment happens in the real world of the classroom (Black, et al, 2004; Furtak & Ruiz-Primo, 2006; Sneider & Wojnowski, 2013; William et al, 2004).

Data from this research confirmed many of the challenges and issues discussed in the literature examining formative assessment. For example, Bennett (2011) claimed that formative assessment is less effective when teachers do not have deep cognitive-domain understanding:

The first implication is that a teacher who has weak cognitive-domain understanding is less likely to know what questions to ask of students, what to look for in their performance, what inferences to make from that performance

about student knowledge, and what actions to take to adjust instruction (Bennett, 2011, p. 15).

Teachers in this study reported uncertainty about their own subject-matter knowledge combined with hesitation and doubt regarding how to address student misconceptions. The sincere desire of each participant to improve their science instruction was evident throughout the entire study, but lack of PCK certainly prevented teachers from using Keeley probes most effectively.

Second, in both Furtak and Ruiz-Primo's (2006) study and the present study, teachers did not frequently use information from student responses to improve student learning. Furtak and Ruiz-Primo found that only about 25% of informal assessment conversations used information from student responses to improve student learning.

Teachers in this study reported difficulty making inferences about student understanding and the steps they might take to remedy misconceptions. Obviously, the teachers made inferences about student understanding throughout classroom inquiries and conversations, but they had difficulty analyzing information from the probes and deciding what might be done to further understanding. Bennett (2011) discussed the inferential nature of formative assessment, because we cannot know with certainty what understanding exists inside student's minds:

We can only make conjectures based on what we observe from such things as class participation, class work, homework, and test performance. Backing for the validity of our conjectures is stronger to the extent we observe reasonable consistency in student behavior across multiple sources, occasions, and contexts. Thus, each teacher-student interaction becomes an opportunity for posing and refining our conjectures, or hypotheses, about what a student knows and can do, where he or she needs to improve, and what might be done to achieve that change (Bennett, 2011, p. 17).

While one of the goals of formative assessment is to validate our conjectures about what students think and understand, identifying what each student needs and what actions to take to improve learning are even more crucial.

This research also relates to the work of Coffey and her colleagues (2011). Teachers reported becoming more aware of student ideas due to the Keeley probe. Coffey (Coffey, et al, 2011) argued that formative assessment should be oriented towards responsiveness to students' ideas and practices in teacher education, as opposed to specific strategies. However, in this study, specific strategies may have helped teachers improve student learning. Teachers in this study showed respect for students as thinkers, but that awareness did not inform their next moves.

Data from this study showed that instruction related to implementation of Keeley probes did not display the key components of formative assessment. According to Black & William (1998b), "Opportunities for students to express their understanding should be designed into any piece of teaching, as this will initiate the interaction through which formative assessment aids learning" (p. 143).

These opportunities would display the following characteristics: (1) the use of classroom discussions, classroom tasks and homework to determine the current state of student learning and understanding with action taken to improve learning and correct mistakes; (2) the provision of descriptive feedback with guidance on how to improve during the learning; and (3) the development of student self- and peer-assessment skills.

Unfortunately, these key components of formative assessment were not illustrated in this research, substantiating the conclusions of both Bennet (2011) and Coffey et al.

(2011) that much work still remains to be done before the promise of formative assessment is realized.

Recommendations

- 1. Develop methods of professional development that help teachers deepen their disciplinary knowledge while improving their skills in analyzing and responding to students' misconceptions.**

Teachers would benefit from methods of individualized professional development in science instruction that would enable them to deepen their content knowledge while developing the analytical skills needed to analyze student responses and make necessary modifications or try new instructional strategies that are most powerful in nurturing true conceptual understanding for students. This recommendation is consistent with suggestions from prior researchers such as Sadler and his colleagues (2013), who suggested using diagnostic identification and remediation of teachers' "knowledge holes" (p. 1043), arguing that professional development programs emphasizing PCK, particularly knowledge of student misconceptions (KOSM) would be more effective. Schulman (1987) argued that teacher education should provide aspiring educators with opportunities to reason their way through effective pedagogy, and other researchers and educational leaders who have attempted to find practical ways to help teachers build pedagogical skill, deep domain understanding, and inferential skills essential to responding to student ideas in a meaningful way (Bennett, 2011; Coffey, et al, 2011; Ruiz-Primo & Furtak, 2006).

Teachers would also benefit from professional development about the use of Keeley probes specifically. Page Keeley and her team provide a variety of workshops at National Science Teacher Association conferences and professional development sessions across the country. Although experienced teachers often try new strategies on their own and set individual goals to improve teaching and learning, greater access to these professional development opportunities would help teachers integrate Keeley probes more effectively into the curriculum.

2. Create an index of Keeley probes aligned to the NGSS.

Teachers have questions about grade-level performance expectations, and how the Keeley probes correlate to NGSS. An index of Keeley probes correlated with each disciplinary core idea specific to each grade level would afford teachers more confidence in addressing grade-level expectations, providing more time for instructional planning. Although the books of Keeley probes include tables of appropriate grade ranges for the concepts addressed by the probes, these tables were developed before the release of the much more specific grade recommendations in the NGSS.

3. Build an activity/resource bank aligned with NGSS and Keeley probes.

Ample evidence in this study suggests teachers became more responsive to student ideas by using Keeley probes. Yet most of the teachers did not notice specific misconceptions expressed by students, nor did they feel confident about what to do to address misconceptions efficiently. Teachers need quick access to useful classroom materials that integrate pedagogical and subject-area knowledge. While the teacher notes provided a few suggestions for instruction, a more comprehensive and targeted bank of

follow-up investigations and activities would provide much-needed instructional support to elementary teachers who struggle to find the most efficient, engaging, and effective visuals, models and investigations to foster deeper scientific understanding.

4. Engage teachers in reflective cycles of implementation.

This study followed teachers through just one cycle of Keeley probe implementation. School leaders would do well to create opportunities for teachers to implement and reflect on new strategies and approaches to improve student learning. For example, Keeley probes might be used to engage teachers in using formative assessment in the classroom, reflect on the process with a coach, then use probes again with further understanding of how to analyze students' misconceptions and how best to address them.

5. Provide day-to-day instructional support for science teachers, especially at the elementary level.

Formative assessment can make a significant impact on teaching and learning when used purposefully and over time. But, according to Keeley, (2008), "Dabbling here and there does not produce significant gains in student learning or teacher performance" (p. 38). Making formative assessment an integrated part of teaching and learning requires substantial changes in school and classroom cultures. Teachers need coaching and support to realize the full potential of formative assessment, and to use Keeley probes most effectively.

Implications for Future Research.

Results from this study point to a variety of future research topics. This research presents data gathered as teachers tried a new strategy for the first time. Future research

which follows teachers through multiple cycles of implementation and reflection would provide a better picture of how teachers use Keeley probes in the classroom, what they learn as they reflect and try this powerful strategy multiple times with a variety of science concepts.

Teachers in this study reported conducting whole-class discussions with students to confront misconceptions and deepen scientific understanding. Further observation and documentation of the configurations and content of these whole-class discussions would provide additional data about how Keeley probes provoke dialogue which improves student learning.

This research supports many of the claims made by Bennett (2011), including the need to develop integrated task sets, projects, diagnostic tests, and observational and interpretive guides to help teachers better integrate the process and methodology of formative assessment with deep domain understanding (p. 19). Further research might examine the process of developing and testing these resources with small groups of reflective, motivated teachers.

Further research is needed to examine how Keeley probes affect student achievement, as well as the most effective strategies to use with the probes. As researchers get a better idea of how teachers use the probes, research data could identify the most efficient, powerful ways to utilize Keeley probes to improve student learning.

Black and Wiliam (1998b) called for living examples of implementation to strengthen the body of literature supporting formative assessment. The research reported in this paper provides one portrait of the work four experienced elementary science

teachers engage in as they attempt to build formative assessment into the curriculum using Keeley probes. More research should examine the practical reality of using formative assessment to improve teaching and learning, in a variety of settings with diverse participants.

Conclusion

Black and his colleagues (2004) posed the question, “What happens when teachers try out ideas which research suggests could benefit their students?” (p. 1). The Keeley probes utilized in this study provoked teachers to become more responsive to student ideas. Coffey and her colleagues (2011) argued that formative assessment should become about engaging with the substance of student ideas and reasoning, responding to those ideas to focus on the disciplinary substance of learning in science. This study shows that, even in such a limited implementation, Keeley probes helped teachers focus on student ideas.

However, teachers need a great deal of time and support to change the nature of classroom interactions, use Keeley probes to respond to student ideas and modify instruction to improve teaching and learning. We know that formative assessment has the potential to improve student learning, and that Keeley probes can be a powerful tool to implement formative assessment. This research illustrates the efforts of a small group of motivated experienced teachers embracing a new strategy, Keeley probes, in their quest to become more responsive to student ideas in science. Significant work remains to support teachers to realize the power of Keeley probes as a formative assessment tool.

References

- American Association for the Advancement of Science (AAAS), (1993).
Benchmarks for science literacy. New York: Oxford University Press.
- Barber, J. (1988). Chemical reactions: Teacher's guide. Berkeley, CA: UC Regents.
- Bennett, R.E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy and Practice*, 18(1), 5-25.
- Bernard, H. Russell (2000). *Social research methods: Qualitative and quantitative approaches*. London ; Thousand Oaks : SAGE. (pp. 443-444)
- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004) Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan* 86(1), 8-21.
- Black, P., & Wiliam, D. (1998a). Assessment and classroom learning. *Assessment in Education*, 5(1), 7-71.
- Black, P., & Wiliam, D. (1998b). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80(2), 139-182.
- Bradley, M. & Harrell, M. (2009). *Data collection methods: Semi-structured interviews and focus groups*. Santa Monica, CA: RAND Corporation.
- Brown, A., & Cocking, R. (2000). How people learn (pp. 285-348). J. D. Bransford (Ed.). Washington, DC: National Academy Press.
- Cochran, K., DeRuiter, J., & King, R. (1993). Pedagogical content knowing: An integrative model for teacher preparation. *Journal of Teacher Education*, 44, 263-272.
- Coffey, J.E., Hammer, D., Levin, D.M., & Grant, T. (2011). The missing disciplinary substance of formative assessment. *Journal of Research in Science Teaching*. 48(10), 1109-1136.
- Driver, R., Suires, A., Rushworth, P., & Wood-Robinson, V. (1994). *Making sense of secondary science: Research into children's ideas*. London and New York: Routledge.

- Dunn, Karee E. and Mulvenon, S. (2009) A critical review of research on formative assessment: the limited scientific evidence of the impact of formative assessment in education. *Practical Assessment Research & Evaluation*. 14(7), 1-11.
- Furtak, E.M., & Ruiz-Primo, M.A. (2006). Informal formative assessment and scientific inquiry: exploring teachers' practices and student learning. *Educational Assessment*. 11(3 & 4), 205-235.
- Furtak, E.M., & Ruiz-Primo, M.A. (2008). Making students' thinking explicit in writing and discussion: an analysis of formative assessment prompts. *Science Education*, 92(5), 799-824.
- Gooding, J. & Metz, B. (2011). From misconceptions to conceptual change. *The Science Teacher* (May/June 2011): 34-37.
- Grounded Theory Institute (2008). What is grounded theory? *GroundedTheory* Institute: a Nonprofit Organization. Retrieved from <http://www.groundedtheory.com> on March 12, 2014.
- Heritage, M., Kim, J., Vendlinski, T., & Herman, J., (2009). From evidence to action: A seamless process in formative assessment? *Educational Measurement: Issues and Practice*. 28(3), 24-31.
- Keeley, P., Eberle, F., and Farrin, L. (2005). *Uncovering student ideas in science, vol. 1: 25 formative assessment probes*. Arlington, VA: NSTA Press.
- Keeley, P., Eberle, F. and Tugel, J. (2007). *Uncovering student ideas in science, vol. 2: 25 more formative assessment probes*. Arlington, VA: NSTA Press.
- Keeley, (2008). *Science formative assessment: 75 practical strategies for linking assessment, instruction, and learning*. Thousand Oaks, CA: Corwin Press and Arlington, VA: NSTA Press.
- Keeley, P. & Sneider, C. (2012). *Uncovering student ideas in astronomy: 45 new formative assessment probes*. Arlington, VA: NSTA Press.
- Koba, S., & Tweed, A. (2009). *Hard-to-teach biology concepts: A framework to deepen student understanding*. NSTA Press.
- National Research Council. (1996). *National science education standards*. Washington, DC: National Academy Press.
- National Research Council (1997). *Science teaching reconsidered: A handbook*.

Washington, DC: National Academies Press.

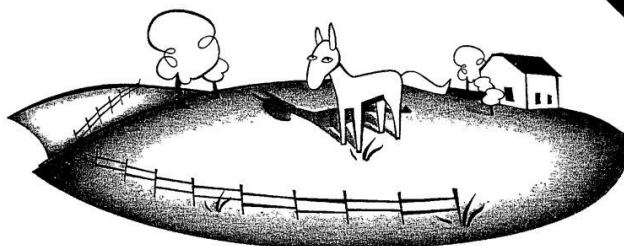
- National Research Council (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: National Academies Press.
- Orland, M., & Anderson, J. (2013) *Assessment for learning: What policymakers should know about formative assessment*. San Francisco: WestEd.
- Osowiecki, A., & Southwick, J. (in press). *Energizing Physics*. New York, NY: Freeman (this course is also described at energizingphysics.com).
- Posner, G., Strike, K., Hewson, P., and Gertzog, W. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education* 66: 211-227.
- Sadler, D. (1998). Formative assessment: revisiting the territory. *Assessment in Education*, 5(1), 77-83.
- Sadler, P., Sonnert, G., Coyle, H., Cook-Smith, N., and Miller, J. (2013) The influence of teachers' knowledge on student learning in middle school physical science classrooms. *American Education Research Journal*, 50(2).
- Sneider, C. & Wojnowski, B. (Eds.) (2013). *Opening the door to physics through formative assessment* (monograph). Portland, OR: Portland State University. Retrieved from National Science Education Leadership website: www.nsela.org/publications.
- William, D., Lee, C., Harrison, C., & Black, P. (2004). Teachers developing assessment for learning: Impact on student achievement. *Assessment in Education: Principles, Policy, & Practice*, 11(1), 49-65.
- Wylie, E. & Heritage, M. (2010). Developing and deepening formative assessment practice. In M. Heritage, *Formative assessment, making it happen in the classroom* (pp. 117-132). Thousand Oaks, CA: Corwin Press.

Appendix A

"Respiration" Probe (Keeley, Eberle, & Dorsey, 2008)

Life, Earth, and Space Science Assessment Probe

Respiration



Put an X next to the organisms on the list that use the process of respiration.

- | | | |
|---|---|---|
| <input checked="" type="checkbox"/> human | <input type="checkbox"/> grass | <input checked="" type="checkbox"/> duck |
| <input checked="" type="checkbox"/> frog eggs | <input type="checkbox"/> mushroom | <input type="checkbox"/> tomato plant |
| <input checked="" type="checkbox"/> fish | <input checked="" type="checkbox"/> chick inside an egg | <input checked="" type="checkbox"/> human body cell |
| <input checked="" type="checkbox"/> germinating seed | <input checked="" type="checkbox"/> horse | <input checked="" type="checkbox"/> apple tree |
| <input checked="" type="checkbox"/> worm | <input type="checkbox"/> single-celled pond organism | <input checked="" type="checkbox"/> honeybee |
| <input type="checkbox"/> bacteria | | |
| <input checked="" type="checkbox"/> butterfly larvae inside a chrysalis | | |

Explain your thinking. How did you decide whether an organism uses the process of respiration?

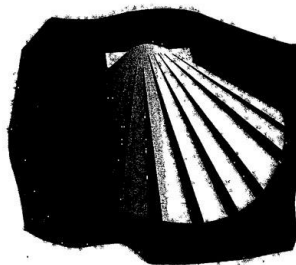
An organism uses the process of respiration by using what is left over from what happens with photosynthesis. That means the organisms gets only the left over energy and the left over sugar from the process of photosynthesis.

Appendix B

"Mountaintop Fossil" (Kelley, Eberle, & Tugel, 2007) Probe

Mountaintop Fossil

The Esposito family went hiking on a tall mountain. Mrs. Esposito picked up a shell fossil on the top of the mountain. The fossil was once a shelled organism that lived in the ocean. The family had different ideas about how the fossil ended up there. This is what they thought:



Mrs. Esposito: A bird picked up the organism and dropped the shell as it flew over the mountain.

Mr. Esposito: Water, ice, or wind eventually carried the fossil to the top of the mountain.

Rosa: A mountain formed in an area that was once covered by ocean.

Sofia: The fossil flowed out of a volcano that rose up from the ocean floor.

Whose idea do you most agree with and why? Describe your ideas about how a fossil could end up on the top of a tall mountain.

I agree with Mr. Esposito, because
it's the most likely idea. If I
made my own idea it would be
the organism climbed up the
mountain and reached the top
but died at the top.

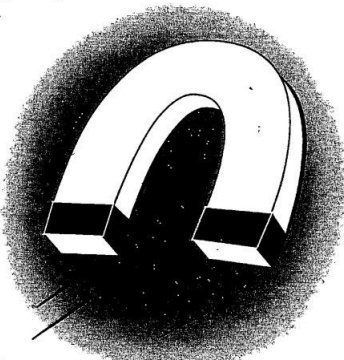
Appendix C

“Is It Matter” (Keeley, Eberle, & Farrin, 2005) Probe

Physical Science Assessment Probes

Is It Matter?

Listed below is a list of things that are considered matter and things that are not considered matter. Put an X next to each of the things that you consider to be matter.

<input type="checkbox"/> rocks	<input type="checkbox"/> salt		
<input type="checkbox"/> baby powder	<input type="checkbox"/> Mars		
<input type="checkbox"/> milk	<input type="checkbox"/> Jupiter		
<input checked="" type="checkbox"/> air	<input checked="" type="checkbox"/> steam		
<input checked="" type="checkbox"/> light	<input type="checkbox"/> rotten apples		
<input checked="" type="checkbox"/> dust	<input checked="" type="checkbox"/> heat		
<input checked="" type="checkbox"/> love	<input checked="" type="checkbox"/> sound waves		
<input type="checkbox"/> cells	<input type="checkbox"/> water		
<input checked="" type="checkbox"/> atoms	<input type="checkbox"/> bacteria		
<input type="checkbox"/> fire	<input checked="" type="checkbox"/> oxygen		<input checked="" type="checkbox"/> gravity
<input checked="" type="checkbox"/> smoke	<input type="checkbox"/> stars	<input checked="" type="checkbox"/> magnetic force	<input type="checkbox"/> electricity

Explain your thinking. Describe the “rule” or reason you used to decide whether something is or is not matter.

These things are probably matter because matter, I think, is something you can't hold.

Uncovering Student Ideas in Science
79

Appendix D*Sample Teacher Responses For Each Conceptual Code*

Code	Example Teacher Response
Teacher pedagogical content knowledge	“I can go to the standards, but that doesn’t always help me. This helps unpack it for me, what are they really asking? What is the underlying concept? That helps me. I don’t have a science background, it’s hard to teach to that conceptual level. I feel like, you are walking around something, but not getting to the real thing.”
Reinforcing big ideas	“This formative assessment probe might be a way to anchor the teaching that I am doing as I go through the box curriculum. I can stray and have conversations to get to the bottom of the big ideas, which doesn’t always happen as I work through a curriculum.”
Learning goals: NGSS and appropriate grade-level expectations	“I hope to find out what they understand about uplift and general plate tectonics—mountain formation. Off the top of my head, I can’t identify (the NGSS performance expectation).” “It’s hard to know what is appropriate grade-level conception, how far to go.”
Modifying instruction	“I am not sure exactly how to go about it, except to have a conversation about what is something common between all these things.”
Student engagement	“The probe led itself to have a discussion right away with them, it seemed to pique their curiosity. I was sort of surprised by that. They were all VERY engaged and interested. They were all engaged in the discussion. ALL of them!”

Finding Time to address misconceptions	“They are still confused about gases, you can’t see it. I wish I could do more follow-up activities.”
Knowing how to address misconceptions	“Do these people have like a connected, like, where to go after this? Because it makes you think about what you are not getting across, but I don’t always know the next really good step, it just needs to be straight-forward and clear, and discovering it versus, you know, “I’m just going to tell you how this works” because it doesn’t work that way. Like, where could I go to get that idea for an activity and discovery instead of just telling them something.”
Student self-awareness	“If I give it before the lesson, will they be paying more attention to the content of the learning? Even if they are not consciously keeping it into their mind, will they be thinking, what am I supposed to be learning? Is this about matter?”
Classroom conversations	“I won’t really give a score—or say right/wrong until the end. Then, later, we will go through and talk about those that are not matter.”
Teacher attitudes towards formative assessment	“Truthfully, I am excited about the probe, getting them to demonstrate their thinking in a different form, so it will be interesting to see what they have to say. I’m curious.”
Multiple forms of instruction	“It helped me hone in on, what exactly is the geology standard at this grade level, how is that a meaningful piece connected to what we are doing, and am I providing enough discussion, examples, visuals, models, as opposed to glossing over with one conversation.”

Appendix E

NGSS Correlation: “Respiration” (Keeley, Eberle, & Dorsey, 2008)

NGSS Standard	Performance Expectation	Disciplinary Core Idea
4-LS1-1*	Construct an argument that plants and animals have internal and external structures that function to support survival, growth, behavior and reproduction	LS1: Plants and animals have both internal and external structures that serve serious functions in growth, survival, behavior, and reproduction
5-PS3-1	Use models to describe that energy in animals/ food (used for body repair, growth, motion, and to maintain body warmth) was once energy from the sun.	PS3.D: The energy released from food was once energy from the sun that was captured by plants in the chemical process that forms plant matter (from air and water). LS1.C: Food provides animals with the materials they need for body repair and growth and the energy they need to maintain body warmth and for motion.
5-LS1-1	Support an argument that plants get the materials they need for growth chiefly from air and water.	LS1.C: Plants acquire their material for growth chiefly from air and water.

Appendix F

NGSS Correlation: “Mountaintop Fossil” (Kelley, Eberle, & Tugel, 2007)

NGSS Standard	Performance Expectation	Disciplinary Core Idea
4-ESS1-1*	Identify evidence from patterns in rock formations and fossils in rock layers to support an explanation for changes in a landscape over time.	ESS1.C: Local, regional, and global patterns of rock formations reveal changes over time due to earth forces, such as earthquakes. The presence and location of certain fossil types indicate the order in which rock layers were formed.
4-ESS2-1*	Make observations and/or measurements to provide evidence of the effects of weathering or the rate of erosion by water, ice, wind, or vegetation.	ESS2.A Rainfall helps to shape the land and affects the types of living things found in a region Water, ice, wind, living organisms, and gravity break rock, soils, and sediments into smaller particles and move them around.
4-ESS2-2*	Analyze and interpret data from maps to describe patterns of Earth’s features.	ESS2. The location of mountain ranges, deep ocean trenches, ocean floor structures, earthquakes, and volcanoes occur in patterns. Most earthquakes and volcanoes occur in bands that are often along the boundaries between continents and oceans. Major mountain chains form inside continents or near their edges. Maps can help locate the different land and water feature areas of Earth.

Appendix G

NGSS Correlation: Is It Matter? (Keeley, Eberle, & Farrin, 2005)

NGSS Standard	Performance Expectation	DCI
5-PS1-1*	Develop a model to describe that matter is made of particles too small to be seen.	PS1.A: Matter of any type can be subdivided into particles that are too small to be seen, but even then matter still exists and can be detected by other means. A model showing that gases are made from matter particles that are too small to see and are moving freely around in space can explain many observations, including the inflation and shape of a balloon and the effects of air on larger particles or objects.
5-PS1-2	Measure and graph quantities to provide evidence that regardless of the type of change that occurs when heating, cooling, or mixing substances, the total weight of matter is conserved.	PS1.B: No matter what reaction or change in properties occurs, the total weight of the substance does not change.
5-PS1-3*	Make observations and measurements to identify materials based on their properties.	PS1.A: Measurements of a variety of properties can be used to identify materials. (mass and weight are not distinguished at this grade level).

5-PS1-4	Conduct an investigation to determine whether the mixing of two or more substances results in new substances.	PS1.B: When two or more different substances are mixed, a new substance with different properties may be formed.
---------	---	--

Appendix H

“Respiration” (Keeley, Eberle, & Dorsey, 2008) Sample Student Responses

No Data	Inaccurate/Naive	Moving Towards Scientific Understanding	Very Close to Accurate Scientific Explanation
I just guessed	I thought whether or not it was a thing that made noise because you need air for that.	Respiration is breathing air so I crossed out everything that breaths air.	I knew that all the producers use the process of respiration so I sed grass, mushroom, tomato plant, and apple tree. I also knew all animals use the process of respiration so I sed human, fish, worm, horse, duck, and honey bee. I don't know if bacteria use this process though.
Well I don't know how to explain	I figured that things inside airtight seals don't use oxygen or air. Life forms that have already been exposed to air.	I know that respiration is the left over things that happen in photosynthesis so I bas it off that.	I know that respiration is the waste from photosynthesis like oxygen that helps plants and animals to live for example a human is a living thing that needs air and plants that depend on humans by using there waste and the plants use that for photosynthesis so plants need the air so all living things need respiration.

No Data	Inaccurate/Naive	Moving Towards Scientific Understanding	Very Close to Accurate Scientific Explanation
	I knew plants process oxygen so that's a no.	I decided by seeing or determining whether or not the use air or oxygen to survive.	

Appendix I**“Mountaintop Fossil” (Kelley, Eberle, & Tugel, 2007) Sample Student Responses**

	Mrs. Esposito	Mr. Esposito	Rosa* (best response)	Sofia
Total = 24	4	5	11	4
Sample explanations	“...It makes more sense that bird would take the fossil up to the mountaintop.”	“...the organism climbed up the mountain and reached the top but died at the top.”	“My idea how the fossil got there is that a mountain formed over lots of years under the water.”	“...Most rock comes from volcanos, and if the volcano was in the ocean then the water would cool the lava into rock on the seashell.”
		“...Water and wind can carry a shell to the top of a mountain. Lava would burn it, and why would a bird drop it?”	“...Sinse it was fossilized it was probably old so it makes sense that a mountain would form there like an island.”	“The bone came out of a volcano that came from the lower part of the ocean.”
		“Maybe a wave came in and had the fossil on it, and when it crashed down and the fossil went on the mountain.”	“A fossil could be dogged up from a settler or it could have gotten wind and it blew from somewhere else.”	“I which a lot of since (science) shows and thare is ever ware bones in the deserts. T is a fan word for shellfish called crusushions. I think that it used to be a reef that rose up from the ocean.”

Appendix J

“Is It Matter” (Keeley, Eberle, & Farrin, 2005)

Jason’s Sample Student Responses

<i>Inaccurate/Naive</i>	<i>Moving Towards Scientific Understanding</i>	<i>Very Close to Accurate Scientific Rule</i>
“...something magnetic”	“something you cannot hold but you know it is there	“If it has mass it is matter.”
“Things you can not touch but they are air.”	“I did a x for every thing that you can see.”	“...something you can touch.”
“...matter, I think, is something you can’t hold.”	“liquids or gas”	“...everything that are made out molecules are matter.”
“Matter is like, something you can’t see or something that is made of a type of energy in it.”	“Something that fills something up.”	“Things that you can hold/things that have weight. Things you can see.”
“...Something that does not take up space.”		
“I’m mainly just guessing.”		

Appendix K

“Is It Matter?” (Keeley, Eberle, & Farrin, 2005).

Christopher’s Sample Student Responses

<i>Inaccurate/Naive</i>	<i>Moving Towards Scientific Understanding</i>	<i>Very Close to Accurate Scientific “Rule”</i>
“...something magnetic”	“something you cannot hold but you know it is there	“If it has mass it is matter.”
“Things you cannot touch but they are air.”	“I did a x for everything that you can see.”	“...something you can touch.”
“...matter, I think, is something you can’t hold.”	“liquids or gas”	“...everything that are made out molecules are matter.”
“Matter is like, something you can’t see or something that is made of a type of energy in it.”	“Something that fills something up.”	“Things that you can hold/things that have weight. Things you can see.”
“...Something that does not take up space.”		