Winter 12-27-2018

Teachers' Professional Knowledge and Formative Assessment Practices: An Empirical Study from Middle School Earth Science Instruction in the Context of an Educational Improved Community

Laura Millay
University of Maine, laura.millay@maine.edu

TEACHERS' PROFESSIONAL KNOWLEDGE AND FORMATIVE ASSESSMENT PRACTICES:

AN EMPIRICAL STUDY FROM MIDDLE SCHOOL EARTH SCIENCE INSTRUCTION

IN THE CONTEXT OF AN EDUCATION IMPROVEMENT COMMUNITY

By

Laura Millay

B.A. Brown University, 2009

A THESIS

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Science

(in Teaching)

The Graduate School

The University of Maine

December 2018

Advisory Committee:

Mitchell Bruce, Professor of Chemistry, Advisor

Shirly Avargil, Assistant Professor of Science and Technology, Technion

Jon Shemwell, Associate Professor of Curriculum and Instruction, University of Alabama

TEACHERS' PROFESSIONAL KNOWLEDGE AND FORMATIVE ASSESSMENT PRACTICES:

AN EMPIRICAL STUDY FROM MIDDLE SCHOOL EARTH SCIENCE INSTRUCTION

IN THE CONTEXT OF AN EDUCATION IMPROVEMENT COMMUNITY

By Laura Millay

Thesis Advisor: Dr. Mitchell Bruce

An Abstract of the Thesis Presented
in Partial Fulfillment of the Requirements for the
Degree of Master of Science
(in Teaching)
December 2018

While there is widespread agreement that effective formative assessment supports student learning in science, the knowledge teachers need in order assess learning remains sparsely studied. In 1999, Magnusson, Krajcik, and Borko (MKB) proposed that Pedagogical Content Knowledge (PCK, Shulman 1986, 1987) is a distinct knowledge domain that synthesizes three base domains that include science subject matter, pedagogy, and context. The MKB model identified Assessment Knowledge as one of five components of PCK. Since 1999, several studies have used the MKB framework, but have left Assessment Knowledge underdefined. In 2012, Avargil, Herscovitz, and Dori proposed a revision based on empirical study, putting Assessment Knowledge outside of and above PCK. This empirical study seeks to clarify the theory and definition of Assessment Knowledge by investigating the knowledge teachers use when planning and carrying out formative assessment in their classrooms. Methods used in this study are grounded theory (Charmaz, 2006), phenomenology (Groenewald, 2004), and case study approaches (Yin, 2017). The data gathered to inform the study include multiple cycles of interviews with four teachers, observations of classroom teaching and assessment, and classroom artifacts including student work. In this thesis, case studies from two formative assessment cycles are presented and discussed.

We found that in the context of teaching with new science materials, teachers relied on diverse knowledge to inform their assessment decisions. Orientations, a component of PCK according to the MKB model, influenced assessment decisions in ways that align with the MKB framework. But other aspects of how teachers' PCK influenced assessment decisions can be better described using a modification to the MKB model that was introduced by Park and Chen in 2012. How professional knowledge shaped assessment decisions differed across cases, showing that interactions among the domains and components of professional knowledge are complex.

In each case, the goals of assessment were for teachers to increase their Knowledge of Students and Knowledge of Instruction, which are both components of PCK according to the MKB framework. But the knowledge gained by teachers through the assessment process differed in the two cases. Lack of alignment across knowledge domains and components constrained learning in one case, while alignment supported learning in the other.

One implication of these findings is that practicing the alignment of assessment tasks can be a pathway for teachers to develop their professional knowledge as they synthesize multiple knowledge domains and components, and test and reflect on their decisions. Another implication is that refinement of current theoretical frameworks may be needed in order to better illustrate the shaping role of orientation as well as the complex influencing relationships among the knowledge domains and components. This study motivates additional case studies to understand factors that shape how knowledge domains and components interact, as well as further investigation of ways to support teachers in developing alignment across the domains and components of their knowledge.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Mitchell Bruce, for supporting and encouraging me throughout the thesis process. I have deeply enjoyed our many conversations and discussions about the topics covered in this thesis, as well as many others. I've very much appreciated the opportunity to be curious, ask questions, and look for answers with a mentor who really cares and who brings tremendous insight and perspective about the topics of investigation as well as the process of pursuing knowledge as a shared endeavor. I would also like to thank my committee members, Dr. Shirly Avargil and Dr. Jonathan Shemwell, for their guidance, ideas, and support throughout this process. I deeply appreciate their willingness to share their expertise and wisdom with me at every point throughout this process, and for as long as I have needed.

Dr. Michael Wittmann and Dr. Francois Amar helped with the early design of this study. Dr. Wittmann also worked with me to analyze data from classroom video and observations and has provided feedback and been a sounding board throughout the project. Paige Gallagher worked with me on this project for two summers, helping with transcription and analysis of classroom artifacts and contributing to many discussions about how to make sense of the interview data. Dr. Natasha Speer has provided insightful comments and good advice at several points in this project, some of which I have been wise enough to follow, and some of which I understand better now.

Dr. Susan McKay initially recruited me to the MST Program and has since provided me and so many others with valuable professional opportunities through the RiSE Center. These have expanded my thinking and experience and have made it possible for me to pursue and continue moving forward with this project. I would also like to thank Dr. Carolina Alvarado, for

pushing me to write my first three chapters and for being constantly enthusiastic, encouraging, and fun to work with. Dr. Benedikt Harrer's doctoral research and his careful approach to research methodology along with his willingness to share his work and ideas supported me during the design of this project.

I would also like to thank the RiSE staff: Marina Van der Eb, Erin Vinson, Beth ByersSmall, Beth Muncey, Maureen Raynes, and Deb Shulman, for being the most wonderful colleagues a person could ask for and for supporting me during this final push to complete my thesis. Also, thanks to the RiSE faculty and graduate students many of whom have contributed to this work with their comments and suggestions. The RiSE Research Group has been a wonderful forum for sharing ideas and getting feedback and has been extremely helpful. And finally, I would like to thank my family and friends, especially Brian Roth, for supporting me in completing this work.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1. Formative Assessment of Science Learning

The practice of formative assessment is at the core of responsive science teaching and can play a key role in supporting the development of teachers' professional knowledge. Over the past two decades, formative assessment has become increasingly integrated into research-based curricular materials and is recognized as an integral part of normal classroom practice for science teachers. The trend toward standards-based assessments and proficiency-based graduation in many states has highlighted the importance of classroom assessment for purposes of demonstrating proficiency and standardizing instruction. A recent National Resource Council publication on aligning assessment to Next Generation Science Standards (Pellegrino, Wilson, Koenig, & Beatty, 2013) highlights the need for ongoing formative assessments that support teaching and learning, in addition to the summative assessments used to demonstrate proficiency. Formative assessment is intended to be an ongoing process that is a part of learning, where teachers gather information about student learning and student difficulties, get feedback about the effectiveness of instruction and information about how to target additional instruction, and give feedback to students. Increasingly, the process is also intended to include students in monitoring and evaluating their own learning.

1.2. Assessment Knowledge

From a theoretical and research standpoint, assessment sits at a crucial intersection of teachers' professional knowledge. Magnusson, Krajcik, and Borko in a 1999 paper on Pedagogical Content Knowledge (PCK) for science teachers place Knowledge of Assessment as one of five areas of professional knowledge: Orientations to Science Teaching, Knowledge of

Student Understandings, Knowledge of Instructional Strategies, Knowledge of Science Curricula, and Knowledge of Assessment. Their definition of Knowledge of Assessment contains two areas of knowledge: knowledge of what to assess (dimensions of science learning to assess) and knowledge of how to assess (methods for assessing science learning).

Building on this construct of Knowledge of Assessment, several other research papers have used the theoretical construct of PCK including Knowledge of Assessment to describe learning processes for pre-service teachers or for teachers as they go through professional development, particularly in teaching a new curriculum. The existing literature suggests that Knowledge of Assessment may play a moderating or synthesizing role in developing other knowledge. Avargil, Herscovitz, and Dori (2012) suggested that Knowledge of Assessment (or Assessment Knowledge, AK) represents a synthesis of teachers' content knowledge and other aspects of their PCK. For example, while designing questions for the purpose of assessing student learning about a new curriculum, teachers drew upon their content knowledge and on their knowledge of curriculum, instruction, and student thinking. This suggests that having teachers develop assessments could be a way to develop other aspects of teachers' knowledge. Other works have suggested that teachers' knowledge of tools and strategies for assessment influences their rate of learning to effectively teach a new curriculum. Heinz, van Driel, and Verloop, for example, found that teachers who had knowledge of more assessment methods developed more knowledge about instructional strategies and their students over the period of the study (2008). The suggested mechanism of this learning was that higher levels of assessment knowledge led to more effective assessment, leading to more useful information about instruction, and leading to teachers' knowledge development. These studies suggest that Knowledge of Assessment is a body of knowledge that supports both teaching practice and teachers' learning over time.

1.3. Purpose of this Study

This study seeks to deepen our understanding of the body of knowledge that a teacher needs in order to skillfully assess student learning of key ideas in science, using appropriate assessment methods. For this study, we have focused on formative assessment, as the practice of assessing for the purpose of informing teaching and learning. The practice of formative assessment as a component of classroom teaching and assessment represents a recent shift from traditional expectations for teachers and is an area requiring additional study (Kolomuç, 2017; Xu & Brown, 2016). In addition, we have focused on the disciplinary substance of formative assessment for teaching science; that is, we focused this study on the assessment specifically of core ideas in science, rather than on many other elements related to teaching and learning that might be assessed in a science classroom, as recommended by Coffey, Hammer, Levin, and Grant (2011). Our work seeks to contribute to an understanding of how knowledge shapes teachers' decisions about formative assessment; specifically, we seek to accomplish the following goals:

1) To understand how knowledge can contribute to middle school teachers' decisions about what to formatively assess in science learning
2) To understand how knowledge can contribute to middle school teachers' selections of methods for formative assessment of science learning

In order to gain understanding of how knowledge contributed to middle school teachers' decisions about what and how to formatively assess science teaching and learning, we developed two case studies of formative assessment cycles in middle school science classrooms. Through interviews conducted before and after instances of formative assessment we explored teachers' decisions about assessment and their reasoning about and reflections on those

decisions. The case studies were then interrogated with regard to a set of research questions that focus on teachers' goals, beliefs, and knowledge base, which are primary elements in Schoenfeld's (1998) model for describing mechanisms underpinning teachers' decisions. The research questions are:

1) How did teachers describe and reflect on their goals about what to assess and the factors they considered to be important in deciding what to assess?
2) How did teachers describe and reflect on their plans about how to assess and the factors they considered to be important in deciding how to assess?
3) How did teachers describe and reflect on what they considered to be important about the student feedback that was elicited through assessment?
4) How did teachers reference their professional knowledge and beliefs while describing and reflecting on the decisions they made about assessment?

While analyzing the data we drew upon a theoretical construct proposed by Magnusson, Krajcik, and Borko (1999) for the domains and elements of professional knowledge that inform teaching practice, and the relationships between the domains and elements. Several empirical studies have used the Magnusson, Krajcik, and Borko model to discuss relationships between the elements of teachers' professional knowledge for teaching, including their knowledge about assessment (Avargil, Herscovitz, & Dori, 2012; Hanuscin, Lee, & Akerson, 2011; Henze, van Driel, & Verloop, 2008). Some studies have considered possible modifications of the Magnusson, Krajcik, and Borko model in order to better describe qualities of teachers' knowledge that have been observed through empirical study (Avargil et al., 2012; Park & Chen, 2012). These will be discussed in detail in the next chapter.

By using the Schoenfeld (1998) and Magnusson, Krajcik, and Borko (1999) models to frame our data analysis, we seek to contribute to a practical understanding of how teachers'

knowledge relates to their decision-making about formative assessment. This work is intended to be helpful in considering the design of both pre-service teacher education and professional learning activities to support current teachers in carrying out formative assessment in their classrooms; an activity that has been shown to be challenging for teachers to enact (Cisterna & Gotwals, 2018). This work is also intended to make a theoretical contribution by providing rich examples of how multiple domains and elements of knowledge can contribute to teachers' thinking about what and how to formatively assess in the science classroom.

1.4. Empirical Methods

This study was conducted in the context of a Mathematics and Science Partnership in which a group of teachers partnered with University faculty and other community partners to select new, research-based curricular materials for teaching science. Data were gathered during the pilot year of teaching the new curriculum over several months from December of 2011 to May of 2012 and in multiple cycles as teachers conducted successive formative assessment activities in their classrooms. The full corpus of data included baseline interviews at the start of the study about past and present assessment practices and the teachers' ideas about assessment, pre-assessment interviews regarding assessment selection, classroom observations, classroom artifacts including student work on written assessments, follow-up interviews with teachers about their uses of the assessment artifacts, and additional cycles as new assessments were chosen, used, and evaluated. Our goal was to assemble a rich set of data that would give us an authentic perspective on the kinds of assessments teachers chose for classroom use, their plans for using the assessments and how those plans were implemented, and the outcomes of assessment. In this thesis, we will address findings from the interview data.

CHAPTER 2

LITERATURE REVIEW

2.1. Classroom Assessment

In a science classroom, student learning may be assessed for multiple purposes, including for the purposes of assigning grades, for purposes of external evaluation (from school administrators or outside entities), and for purposes of informing instruction. At times, a single assessment may serve multiple purposes. Multiple distinctions are made in literature on assessment based on the purposes of assessment ("formative" and "summative" as distinguished in Atkin, Black, & Coffey, 2001); location of assessment ("internal" or classroom assessments and "external" assessments typically used for accountability purposes as described in Pellegrino, Wilson, Koenig, & Beatty, 2013, Chapter 4) and timing of assessment (e.g., "proximal formative assessment" as assessment conducted in the moment of teaching described in Erickson, 2007).

2.2. Formative Assessment

This study focuses on the teacher knowledge underpinning formative assessment practices. Formative assessment is assessment that informs instruction. From Black et al., "An assessment activity can help learning if it provides information to be used as feedback, by teachers, and by their students, in assessing themselves and each other, to modify the teaching and learning activities in which they are engaged. Such assessment becomes "formative assessment" when the evidence is actually used to adapt the teaching to meet learning needs."(Black, Harrison, Lee, Marshall, & Wiliam, 2003) Formative assessment therefore includes at least two crucial components: (1) an assessment activity that provides teacher and students with (2) information that informs instruction.

Black and Wiliam (2009) further defined "formative assessment" as the following: "Practice in the classroom is formative to the extent that evidence about student achievement is elicited, interpreted, and used by teachers, learners, or their peers, to make decisions about the next steps in instruction that are likely to be better, or better founded, than they would have taken in the absence of the evidence that was elicited."(Black & Wiliam, 2009) This definition makes specific reference to assessment being defined as formative when evidence is gathered that informs instruction by influencing the decisions made by teachers, learners, or others involved in the teaching and learning process.

Black and Wiliam (2009) go on to discuss the decision-making as an integral component of formative assessment, stating that "formative assessment is concerned with the creation of, and capitalization upon, 'moments of contingency' in instruction for the purpose of the regulation of learning processes."(2009, p.10) Moments of contingency are points at which decisions are made about the direction of instruction. These moments could be synchronous or asynchronous with instruction. A synchronous moment of contingency would be simultaneous to instruction, as with a real-time decision made during classroom discussion. For example, a teacher might decide to ask a particular question or to provide a specific example during a classroom discussion in response to an idea raised during the discussion. An asynchronous moment of contingency occurs when there is a gap between instruction and the point at which a decision is made about instruction. For example, a teacher might see something while correcting homework that would influence planning for the next lesson, or might consider year-end test results when planning for a future year. (Black & Wiliam, 2009)

2.3. Formative Assessment of Science Learning

Cowie and Bell (1999) modeled formative assessment based on 2 years of data collection with 10 science teachers. Their model included data gathered from two categories of formative assessment that they defined as "planned" and "informal" assessment. Planned assessment (as with a written assessment task) involved a process of eliciting, interpreting, and acting on information around an assessment purpose. Interactive assessment (as with assessment during classroom discussion) involved a process of noticing, recognizing, and responding to information around an assessment purpose and occurred on a faster timeline than planned assessment. According to Cowie and Bell, both types of formative assessment relied on teachers' professional knowledge (Shulman, 1986, 1987). For example, as stated in Cowie and Bell regarding the step of "interpreting" a planned formative assessment: "The teachers indicated that their knowledge bases (Shulman, 1987) were important factors in in their being able to interpret the information they had collected in their planned formative assessment. They indicated that interpreting involved using their content knowledge, general and content pedagogical knowledge, curriculum knowledge of learners and their students in particular, a knowledge of educational contexts and a knowledge of educational aims and goals."(Cowie & Bell, 1999, p.105) They do not, however, go into detail regarding the relationships between teachers' professional knowledge base and their formative assessment practices.

In a critique of research on formative assessment practices in science classrooms, Coffey, Hammer, Levine, and Grant (2011) called on the research community to re-examine the depth at which we study formative assessment. They argued that the quality of formative assessment should be judged by the extent to which formative assessment informs and thereby deepens the teaching and learning of core science content (Coffey et al., 2011). According to

their work, exemplar formative assessment practices from prior publications fall short of the kinds of assessment that would be most fruitful for guiding science instruction. This critique raises additional considerations regarding the knowledge teachers and pre-service teachers need in order to assess students' thinking and learning about core science content.

2.4. Teachers' Professional Knowledge for Teaching

In 1986 and 1987, Lee Shulman presented the education research community with the idea that teachers have and/or develop knowledge unique to the profession of teaching and to the specifics of the subject matter that they teach (Shulman, 1986, 1987). He also began to define the domain of Pedagogical Content Knowledge (PCK) as knowledge that guided skillful teaching decisions regarding how to teach specific content to a group of students in a specific context. A large body of work has built upon and fleshed out multiple constructs for teachers' professional knowledge for teaching (Carlsen, 1999; Grossman, 1990), as well as knowledge for teaching science (Abell, 2007; Magnusson, Krajcik, & Borko, 1999; Park & Oliver, 2008) and mathematics (Ball et al., 2000; Hill, Ball, & Schilling, 2008). The multiple interpretations and representations of teachers' knowledge have sparked a call to develop a consensus model of knowledge for teaching, and a book on this subject is due to come out in January, 2019.

Efforts to shift the conversation about professional knowledge beyond defining the components of teachers' professional knowledge for teaching and toward suggesting mechanisms for how teachers make decisions in context includes work by Schoenfeld (1998). Schoenfeld proposed a model for understanding and predicting teachers' decisions based on seeking to understanding multiple elements of teachers' thinking, including goals, beliefs, and professional knowledge as well as action plans, ways in which professional knowledge is

organized and accessed, and how teachers envision contingencies and possibilities related to instruction.

2.5. Teachers' Professional Knowledge for Assessment

In 1999, Magnusson, Krajcik and Borko modeled Pedagogical Content Knowledge for Science Teaching and specifically included Knowledge of Assessment as a component in the model. Other elements of PCK included in this model include orientation to science teaching, knowledge of science curricula, knowledge of students' understanding of science, knowledge of instructional strategies, and knowledge of assessment of science literacy. According to this model, Knowledge of Assessment is defined as "a teacher's professional knowledge of what to assess and how to assess in the context of a science classroom."(Magnusson et al., 1999) This definition breaks into two components: knowledge of dimensions of science learning to assess, and knowledge of methods for assessing science learning.

In her 2007 review of research on teacher knowledge, Sandra Abell depicted a broader range of knowledge and situated PCK within a conception of an overall knowledge base for science teachers. In her review, she lends part of a page to the idea of assessment knowledge, relying on the Magnusson, Krajcik, and Borko definition of knowledge of domains and methods of assessment (again without any values added for what should be assessed.) She cites a handful of empirical studies, but states that, "More studies are needed to better understand what teachers know about assessment, and how they design, enact, and score assessments in their science classes."(Abell, 2007)

Since then, multiple frameworks for knowledge of assessment have been proposed. Pellegrino and colleagues (National Research Council, 2001) proposed an assessment triangle, with cognition, observation, and interpretation as the three vertices. Cognition referred to

teachers' beliefs and understandings about how students think and learn. Observation referred to the specifications for selecting or designing assessment tasks that would elicit information from students that would be useful to the purposes of assessment. Interpretation referred to methods and tools used to reason about the observations made through assessment.

Abell and Siegel (2011) extended elements of both this and the Magnusson, Krajcik, and Borko model to propose a broader framework for what teachers need to know in order to be assessment literate. Through gathering and coding empirical data according to Magnusson, Krajcik, and Borko's definition of Knowledge of Assessment, Abell and Siegel identified additional areas that teachers considered, beyond what and how to assess. These included knowledge about purposes of assessment and knowledge of how to interpret findings from assessment.

Xu and Brown (2016) reviewed literature focused on assessment literacy from 1985 to 2015 and proposed a reconceptualization of the assessment literacy (AL) framework. In their model, teachers' knowledge formed the base of a pyramid that then built up teachers' conceptions of assessment, assessment literacy in practice, and identity as assessors. Their model sought to contribute, "to the theorization of AL by moving the field beyond a focus on the knowledge base to consideration of a situated, dynamic, and evolving system in which teachers constantly make compromises among competing tensions…"(Xu & Brown, 2016, p.159). They also called for additional empirical studies to investigate the framework as well as the knowledge base needed for assessment, with one area for further study being to gain an understanding of how contextualized the knowledge base needs to be.

2.6. Empirical Studies of Knowledge of Assessment

Since 1999, multiple studies have suggested that Assessment Knowledge plays a critical role in development of other aspects of teachers professional knowledge (Goodnough & Hung, 2009; Hanuscin et al., 2011; Henze et al., 2008; Kaya, 2008). One mechanism by which Assessment Knowledge may influence other aspects of professional knowledge is that the skill with which teachers conduct classroom assessments shapes the kinds of feedback they receive, both about their students' knowledge and learning and about their own instructional strategies. In addition, designing assessments has been suggested as a method for conducting teacher professional development that builds teachers' PCK, and their Content or Subject Matter Knowledge, because in designing and evaluating assessments, teachers must build on their knowledge of the content, their knowledge of students and curriculum, and on other aspects of their PCK (Avargil et al., 2012). Assessment Knowledge may therefore play a critical role not only in teachers' assessment practices, but also in their instructional practices and PCK development over time.

The next four studies reviewed have used Magnusson, Krajcik, and Borko's (1999) model of PCK as a framework to guide aspects of study design, to lend order to empirical observations, and to help explain findings. For example, a study by Henze, van Driel, and Verloop (Henze et al., 2008) followed nine experienced teachers through their first three years of teaching the topic of "modeling the solar system" in a reform curriculum that emphasized aspects of inquiry and Nature of Science (NOS). Through an interview, repeated each year for three years with each teacher, the authors investigated four aspects of the teachers' PCK within the particular topic. These four aspects, modeled after Magnusson et. al.'s framework, were: knowledge about instructional strategies; knowledge about students' understanding; knowledge about assessment of students; and knowledge about goals and objectives of the topic in the

curriculum. After the initial interview, teachers were characterized as having either Type A PCK or Type B PCK. Type A PCK involved an orientation toward science teaching that was relatively teacher-centered with a focus on content. Type B PCK involved an orientation toward science teaching that was more student-centered with a focus on model content, model production, and thinking about the nature of models.

For teachers with both Type A and Type B PCK, the authors found that their knowledge of goals informed their knowledge of instructional strategies, and that the knowledge of goals did not change over the period of the study. However, there were differences between teachers with the different types of PCK in terms of change over time. In particular, teachers with Type A PCK seemed to increase their knowledge of instructional strategies somewhat, through their increasing knowledge of students' understanding over time. Teachers with Type B PCK, on the other hand, developed increased knowledge in assessment, students' understandings, and instruction. For these teachers, the three components of PCK were linked, with knowledge of instructional strategies informing knowledge of assessment, and knowledge of assessment informing and being informed by knowledge of students' understandings. In addition, there was feedback between teachers' knowledge of their students' understanding and their knowledge of instructional strategies. Teachers with Type B PCK developed more aspects of PCK over time than teachers with Type A PCK.

In a study by Kaya (2008) Turkish pre-service teachers assessment knowledge was categorized as either low or high as they prepared to teach within a national reform-oriented STS curriculum. Pre-service teachers who stated an intention to use "traditional" assessment methods (rather than ones more suited to the reform curriculum) were categorized as having low levels of assessment knowledge relative to the goals of the reform. In order to appropriately implement reforms, the authors stated that pre-service teachers should be prepared to use a

variety of forms of assessment to support student learning, including concept maps and portfolios, rather than using tests for the purpose of measuring student learning.

Kaya's study uses a content test covering five aspects of ozone-layer depletion and an interview with selected participants in order to map quantitative links between Magnusson, Krajcik, and Borko's five elements of PCK, and the pre-service teachers' content knowledge. In all categories other than assessment knowledge, pre-service teachers with higher levels of content knowledge also demonstrated higher levels of PCK within the topic. Only a small percentage of students demonstrated knowledge of assessment that would be in-line with the reform curriculum and no statistically significant correlation was found between content knowledge and assessment knowledge.

A study by Goodnough and Hung (2009) investigated development of PCK for five teachers as they adopted a new problem-based learning approach in their classrooms. Using the Magnussen, Krajcik, and Borko (1999) framework, the study reported that teachers used a variety of assessment methods to assess the different types of learning that students were doing with problem-based learning, and even began having students do self-assessment. The authors state that, "The teachers recognized the integral relationship between assessment and instruction, considering both in an integrated manner during both the planning and implementation stages" (Goodnough & Hung, 2009, p.238).

In another study, Hanuscin, Lee, and Akerson (Hanuscin et al., 2011) used the Magnussen, Krajcki, and Borko (1999) model of PCK to frame a study of PCK in teachers who had undergone 3 years of professional development on teaching about the Nature of Science (NOS). The authors found that the teachers, while successfully improving their students' understandings of NOS, lacked knowledge of ways to assess students' NOS learning. During the study, the researchers measured student learning of NOS and found that the teachers relied on

the information provided by researchers and on their general "sense" about what worked in their own teaching in shaping their views about the success of their instruction. Although the teachers had general assessment knowledge that they used for assessing other topics, they did not know what to assess within NOS. Or as stated by the authors, "…while they had more general knowledge of assessment that they drew on to assess students' ideas about other content, they had not developed topic-specific strategies for assessing students' ideas about NOS" (Hanuscin et al., 2011, p.162). Because of their lack of assessment knowledge for NOS and therefore of assessment of their NOS instruction, teachers had the sense that their students were generally "making connections" (p. 162) about the NOS content they had learned, but could not specifically identify difficulties that their students were having, or strategies that they as teachers used to overcome those difficulties. The authors argue that this constitutes an example of uneven PCK development, in which teachers developed new knowledge of instructional strategies for NOS, but did not develop new assessment knowledge. They go on to describe assessment as necessary feedback for the teacher, for adapting ongoing instruction to the situation and students or for adapting their knowledge for future instruction of the same topic. Without adequate assessment, teachers did not know exactly what their students had learned, where they had trouble, or where instruction may or may not have been effective. This lack of knowledge produced by assessment hampered the teachers' reflective practice and ability to improve their practice for the future. The authors state that, "it stands to reason that teachers' knowledge of assessment might serve as a limiting factor in developing their knowledge of learners. Given the integrated nature of PCK, insufficient development of one component knowledge base can have consequences for enactment of teachers' PCK. For example, without well-developed knowledge of assessment of NOS, the three teachers lack an important source of feedback to inform their instruction."(p. 163)

The authors also suggest a possible connection between assessment and teachers' beliefs. Without accurate assessment, teachers may believe that they are effectively teaching something (like NOS) when they are not. Without accurate assessment, "Teachers' beliefs were based on their own actions (i.e., what they did that they considered to constitute teaching NOS) rather than assessment of the impact of their instruction on students' learning." (p.163) They go on to state that, "equipping teachers with the necessary knowledge of assessment to close the feedback loop between teaching and learning can contribute to further development of their PCK" (p. 165).

2.7. Knowledge for Assessing Science Learning

This study applied Magnusson, Krajcik, and Borko's framework for professional knowledge using methodology that is similar to the work described by Abell and Siegell (2011) in order to map ways in which teachers referenced the multiple elements of their knowledge base while describing and reflecting on decisions about formative assessment. Our methods drew on ideas from Schoenfeld (1998) to conceptualize how teachers' knowledge base (specifically, the knowledge inventory) may relate to their contextualized assessment decisions by probing teachers' goals and beliefs about assessment through interview questions. Our understanding of the nature of assessment knowledge was shaped in part by the work of Avargil, Herscovitz, and Dori (2012) and the finding that teachers synthesize other knowledge elements while designing assessments. We therefore anticipated that teachers would reference multiple elements of their knowledge while describing and reflecting on their assessment decisions. Our methods for conducting this empirical study were tailored in order to provide opportunities for insight into the goals, beliefs, and knowledge that teachers considered while making decisions about formative assessment in the context of middle school Earth science instruction. Through

case studies, we sought to contribute rich empirical examples of how teachers' professional knowledge may inform their formative assessment practices. This work can contribute to the need identified by Xu and Brown (2016) for empirical studies to enhance our understanding of the contextual nature of the professional knowledge that informs assessment practices.

CHAPTER 3

STUDY DESIGN AND DATA COLLECTION

3.1. Context and Study Design

This study adapted a design suggested by Avargil, Herscovitz, and Dori, in order to understand some of the ways in which content knowledge and pedagogical content knowledge intersect as teachers made formative assessment decisions as part of their classroom teaching practice. Avargil, Herscovitz, and Dori (2012) recommended studying teachers' self-designed assessments as a fruitful way to learn about their professional knowledge based on a finding that the task of assessment design required teachers to synthesize multiple aspects of their professional knowledge, including content knowledge and pedagogical content knowledge.

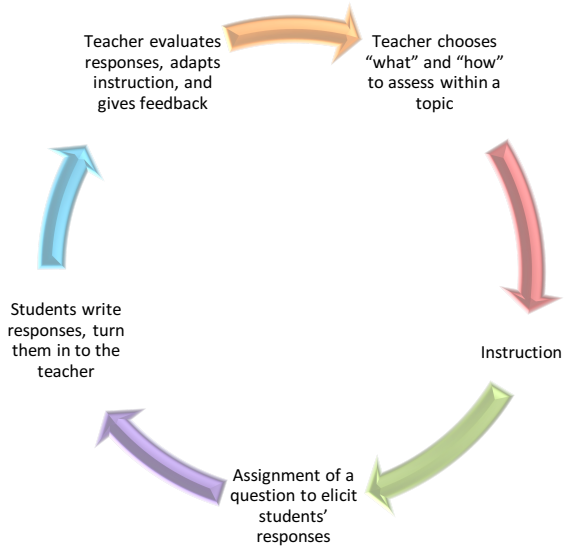
Our study was conducted in the context of an NSF-funded rural Mathematics and Science Partnership (MSP) between University faculty, grade 6-9 teachers, and rural school districts implementing research-based and evidence-based reforms in public school science classrooms. One goal of the MSP was to provide uniform, high-quality middle and high school science curricular materials (including teacher guides, books for students, structured hands-on activities and all of the science materials needed to complete the activities) to more than 100 teachers in 40 schools across multiple school districts in a local control state. Science curricular materials were selected through a year-long task force process, in which middle and high school teachers, University faculty, MSP staff and post-docs, pre-service teachers, and other community partners investigated and vetted curricular materials offered by multiple publishers. Criteria from AAAS (McNeely, 1997) were used in the selection process. One element of the selection criteria was the strength of NGSS-aligned formative assessment practices and tasks that were embedded within the curricular materials through the student books and teacher guides.

Ultimately, three sets of curricular materials were chosen including the Science Education for Public Understanding (SEPUP) Issues and Earth Sciences Curriculum which provides the curriculum context for this study. A small group of teachers were then asked to pilot the materials in their classrooms for the school year following the conclusion of the task force process (the 2011-12 school year). In order to continue evaluation of the curricular materials chosen by the task force, teachers were asked during the pilot year to implement the curricular materials with as much fidelity of implementation as possible. It was therefore undesirable during the pilot year (the year in which this study was conducted) to require teachers to design their own new assessments as suggested by Avargil, Herscovitz, and Dori. Our study design was therefore re-envisioned with the goal of studying teachers' professional knowledge through observations and interviews conducted at key points in their processes of selecting and implementing or formative assessment tasks that the teachers chose, either from the existing palette of assessments embedded within the SEPUP curricular materials, or ones that the teachers designed as a result of their own sense of necessity. Because these curricular materials and the accompanying assessment tasks were new to the teachers and students, we reasoned that we would be able to study teachers' professional knowledge by focusing on the decisions they made about which formative assessment tasks to emphasize, how to have students approach the tasks, and how to score or comment on students' responses. Each assessment decision made in the context of the new materials would represent a moment of contingency (Black & Wiliam, 2009) in which teachers' professional knowledge would play a guiding role. By exploring the decisions teachers made about assessment and the outcomes of those decisions, we expected to find evidence of their professional knowledge.

This study was designed to elicit information about decisions teachers made in the context of planning and implementing assessment activities, evaluating information about

student learning, and designing instructional responses. We considered these activities to constitute a formative assessment cycle that included each essential element of formative assessment. As part of our study design, we made an open-ended request that teachers either select or design a planned assessment activity that would generate some tangible classroom artifact and would be useful to their instruction. The assessments used in the two case studies described in the chapters of this thesis involved Analysis Questions used directly from the SEPUP materials. Figure 1 depicts a possible model for a formative assessment cycle, specifically envisioned for the type of formative assessment task that became the focus for this study. We do not suggest that the process of formative assessment in a typical classroom would always be conducted in this order or would always include all elements. Rather, we suggest that the elements and order depicted here represent one rendering of the core elements of formative assessment as defined by Black and Wiliam (2009).

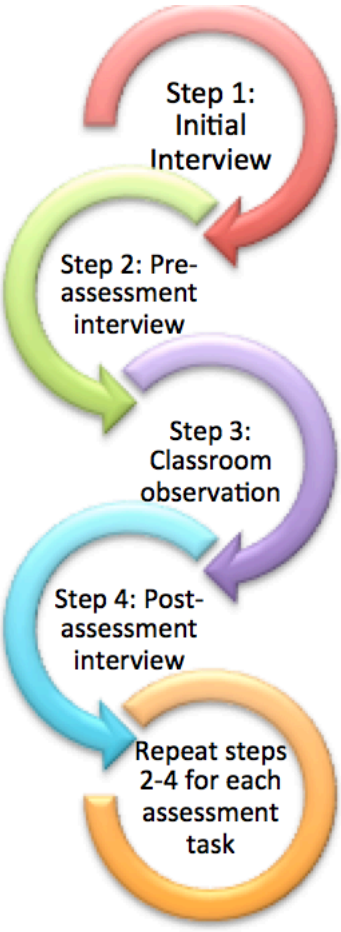
Figure 1. Formative Assessment Cycle



3.2. Participant Selection and Recruitment

In the Fall of 2011 during the pilot year for SEPUP Issues and Earth Science, four teachers were recruited from a 12-teacher cohort of middle school science teachers who had agreed to pilot the new materials. One constraint on participant selection was that the author and primary researcher for this study was assigned through the Partnership to be a "teaching partner" or assistant to the teacher in classrooms of 3 of the 12 teachers. The research team decided that the roles of teaching partner and of researcher should be kept separate, so the three partnered teachers were not considered during the selection process.

Figure 2. Data Gathered at Multiple Points in the Assessment/Instruction Cycle



Four teachers were selected from the larger group of 12, such that the study would gather data from teachers who were diverse in their years of teaching experience and in their level of participation in the MSP as of the time of the study. The four teachers selected for the study had between five and twenty-four years of teaching experience and ranged from most involved in the MSP to least involved in the MSP, of the 12 teachers participating in pilot year. Two of the teachers taught science in the same rural middle school, while the other two teachers taught in two other rural schools. When data collection for this study began, the teachers were nearly 3 months into the pilot year of teaching the new curriculum, and all but one of the teachers had participated in some professional development activities

through the work of the Partnership. A recruitment email was sent three months into the 2011-12 school year. All four of the teachers who were asked to participate agreed to do so.

3.3. Data Collection Overview

Data collection was designed to provide a broad view into teachers' thinking about assessment and instruction both in general and in the context of the new materials they were engaged in piloting as well as the partnership in which they were involved. For each teacher, our data collection included the following elements that will be described in more detail below. These were: (1) a baseline interview followed by a request to each individual teacher to select an assessment task for classroom use and as the focus for our study; (2) a pre-instruction interview about the selected assessment task conducted prior to the assessment task being assigned to students; (3) observation of the classroom instruction related to the assessment task and collection of copies of student work on the assessment task; and (4) a post-instruction interview about the outcomes of the assessment task, conducted after the teacher had evaluated student responses to the assessment task. Steps 2-4 were repeated with some of the teachers, in order to study multiple assessment cycles. (See Figure 2) In addition, each teacher participated in a brief final interview at the end of the 2011-12 school year. A summary of the data collected is provided in Table 1. Interview protocols are provided in Appendices A-C.

Table 1. Summary of Data Gathered

Teacher	Interview 1 ≈ 1 hr each (Baseline)	Interview 2 ≈ 30-60min (Pre-Instruction)	Classroom Observation and Audio 1-3 class periods per cycle	Interview 3 ≈ 60min (Post-Instruction)	Student Work	Interview 4 ≈ 15min (Reflection)
Teacher A	12/12/11	3 cycles	3 cycles	3 cycles	3 completed assessments, 1 Unit Test	6/12
Teacher B	12/15/11	2 cycles	2 cycles +	2 cycles +	2 completed assessments, 1 Unit Test	6/12
Teacher C	1/18/11	-	-	-	-	6/12
Teacher D	2/7/12	1 cycle	1 cycle	1 cycle	i-movies, video of student presentations	6/12

3.3.1. Data Collection Details

A semi-structured baseline interview was conducted with each of the four teachers to discuss past and current assessment practices, ideas and beliefs about purposes and methods of assessment, and the teachers' science and teaching backgrounds. For the baseline interview protocol, see Appendix B. After the baseline interview, participating teachers were asked to choose a formative assessment from within the SEPUP book or to modify/design one of their own for use in their classroom. It was up to the teachers to decide what would be most valuable to their classroom and what they wanted to show to the research team. Given the context described earlier in this chapter regarding an emphasis on fidelity of implementation, both

teachers and researchers were aware that modifications to the curriculum were to be made only if necessary[1].

Teachers were asked to select for research purposes an assessment tasks that exemplified their normal classroom use of formative assessment and that would be useful to their classroom instruction during the pilot year. See Appendix A for the request provided to the teachers. Each teacher individually selected an initial assessment task as the focus for further study and then informed the research team of their choice. The research team had immediate access to most of the assessment tasks, as they were embedded within the SEPUP materials as Assessment Questions that students were to respond to after completing an activity. For any teacher-designed or modified assessments, teachers supplied (via email or in person) copies of the designed or modified versions. Once each teacher chose an assessment task, a semi-structured individual interview was conducted at some point prior to the start of classroom instruction related to the topic to be assessed. In some cases, these pre-instruction interviews were conducted a couple of weeks before the related instruction.

In the pre-instruction semi-structured interviews, teachers responded to questions about their reasons for selecting the assessment task both as a tool for classroom instruction and as a focus for the research. The pre-instruction interview protocol can be found as Appendix C. They described their plans for instruction related to the assessment tasks, their goals for instruction, and their goals regarding the types of information they hoped to obtain through use of the assessment. Teachers were asked to predict what their students would do in response to

[1] "Curriculum" is used here to refer to the collection of materials and to the scope and sequence and instructional strategies recommended by the Partnership. The partnership recommendations to teachers included both published materials and additional materials/modifications/recommendations made by teachers, graduate students, undergraduate students, faculty, and staff during Summer 2011 in preparation for the pilot year. What teachers were piloting was therefore a modified version of SEPUP Issues and Earth Science including some specific recommendations regarding instruction and assessment.

the assessment task including where they might have difficulties, and they were asked to outline their plans for how they might respond to information obtained from students using the assessment task. Teachers were also asked to identify the instruction that was most directly related to the assessment task so that researchers could attend class to observe the related instruction.

In general, the assessment tasks selected by teachers were Analysis Questions found in the SEPUP student book at the end of each activity. Teachers chose analysis questions from a range of activities across multiple units. In addition, there were a few assessments that deviated from the typical pattern, including one end-of-unit assessment designed by a team of teachers and one assessment activity carried over from instruction a teacher had done in previous years. In the cases where embedded analysis questions were chosen, the most relevant classroom instruction related to each question spanned between one and three 45-minute class periods. In the other two cases described, related instruction spanned a longer period of time and it was not possible to observe all of the directly relevant classroom activity.

Classroom observations were conducted by one or two members of the research team. Audio recordings and field notes were taken during each class period that was observed and where possible, video recordings were made as well. In general, the relevant assessment tasks were assigned as in-class activities that were completed while the researchers were in the room. In all, more than 15 separate class periods were observed.

Copies of the students' written work in response to the assessment tasks were made by teachers or the researcher after the teacher had a chance to review and comment, if they chose to provide written comments to students.

After the teacher had an opportunity to review students' responses and to respond in whatever way they chose to the information gathered using the assessment task, another hour-

long semi-structured interview (the post-instruction interview) was conducted. The full post-instruction interview protocol an be found in Appendix D. In this interview, teachers reflected on instruction, on student learning, on the information obtained through use of the assessment tasks, and on any responses they had made to that information (e.g., written comments, verbal feedback, modifications to instruction, etc.) In addition, the researchers probed deeper into anything noticed by the research team during the assessment cycle that might provide insight into teachers' decision-making, including aspects of instruction, of the directions teachers gave when assigning the assessment task, or of the feedback or other responses provided by teachers.

With two of the teachers, the entire data collection cycle including assessment task selection, pre-instruction interview, classroom observation and collection of student work, and post-instruction interview, was repeated. The cycle was repeated twice (for two assessment tasks) selected by one of the teachers and three times (for three assessment tasks) with another. In addition, there were some anomalous opportunities for data collection including one instance in which a member of the research team observed and recording an assessment planning session conducted by the team of two teachers who worked together at a single school.

At the end of the school year, brief semi-structured reflection interviews lasting approximately 15 minutes were conducted with each teacher. Teachers were asked to reflect upon the process of participating in the research as well as on any changes in their overall thinking about assessment as a result of their participation.

3.4. Relationships Between Data Gathered and Research Questions

Table 2 (below) outlines the research questions for this study and how each aspect of data collection relates to each question.

Table 2. Research Questions and Data Used to Inform the Questions

Research Questions	Data Gathered
1. How did teachers describe and reflect on their goals about what to assess and the factors they considered to be important in deciding what to assess?	a. Baseline interviews were conducted and transcribed to provide context by gathering information about teachers' thinking about instruction and assessment in general b. During pre- instruction interviews, teachers were asked to describe their goals for the assessment and what they believed was important to assess, using questions including "what are your goals for the assessment?" and "what do you consider to be important for your students to do with this assessment?" and "what kinds of thinking are you hoping to assess?" Teachers tended to talk about their goals for assessment even earlier in the interview when asked to describe their plans for assessing students. c. During the post-instruction interviews, teachers were asked to reflect on whether they considered the assessment to have been successful in meeting their goals, using questions such as, "what were your goals for this assessment? In what ways were your goals met or not met?" and "what aspects of student thinking did this assess?"

Table 2 continued

2. How did teachers describe and reflect on their plans about how to assess and the factors they considered to be important in deciding how to assess?	a. Baseline interviews asked teachers about their assessment practices and whether they were assessing students differently using SEPUP than they had with their previous instructional materials. Teachers were also asked to discuss an example of an assessment they had used recently. b. In the pre-instruction interviews, teachers were asked about their assessment plans using questions such as, "What will your students do with the assessment?" Teachers tended to describe their planned methods for assessment even earlier in the interviews, when asked in general about their plans for the assessment. c. In the post-instruction interviews, teachers were asked to reflect on what they learned about students' thinking and on the quality of the assessment, using questions such as, "How would you rate this assessment (compared to others in the curriculum or others you have done in the past) – how good of an assessment is it?"
3. How did teachers describe and reflect on what they considered to be important about the student feedback that was elicited through assessment?	a. During the pre-instruction interviews, teachers were asked about what they intended to learn from the assessment, using questions such as, "How will you evaluate student responses?" and "What are you hoping to learn from student responses?". b. During the post-instruction interviews, teachers were asked questions designed to elicit their reflections on the quality of the student feedback elicited through assessing using questions such as, "What was the overall quality of your students' work?" and "Was there anything that stood out for you about their work on this assessment?".
4. How did teachers reference their professional knowledge and beliefs while describing and reflecting on the decisions they made about assessment?	Teachers referenced their knowledge and beliefs throughout the baseline, pre- and post-instruction interviews. Interviews did not specifically ask about teachers' knowledge or beliefs, but were designed to elicit teachers' thinking and reasoning about assessment.

3.5. Research in Partnership

Given that our research was conducted in the context of a Partnership that included a broad array of projects related to research and implementation of professional development and instructional reforms, we also sought to have teachers be partners in this research project. To that end, we did our best to be open about what we were studying and about what we thought we knew or did not know about classroom formative assessment at the middle school level and how teachers might approach it.

CHAPTER 4

DATA ANALYSIS METHODS

4.1. Case Studies

The case studies were assembled from multiple sources of data gathered in relation to the assessment practices of individual teachers. For the purposes of this thesis, each assessment cycle with each individual teacher is a case of planning for, enacting, and evaluating formative assessment. We originally intended to assemble multiple case studies of assessment practice for each teacher in order to develop case studies of each teacher. But while we gathered data in order to accomplish that task, the scope of data analysis and synthesis required to complete these case studies of individuals and their growth over time, is beyond the scope of this thesis. Therefore, the case studies used here are of single assessment cycles.

4.2. Phenomenology

Phenomenology was used in order to make holistic sense of the data from the perspective of the teachers who were the subjects of this research. The author conducted the initial interviews and classroom observations and actively tried to understand the thoughts and motivations of the teachers as they enacted planning, classroom instruction, and evaluation of the assessment. After the assessment cycles were complete, the author listened many times to the recorded interviews and portions of the classroom instruction, reflecting on the meanings of the events and decisions made for the teacher, considering the experiences of students, and reviewing all elements of the interview transcripts. The assessment cycles were discussed at length with research colleagues and portions of the cycles were used to contribute to talks and a poster that focused on the decisions teachers made and the assessment activities they engaged in, and the meanings those decisions and activities held for the participants. In addition,

reflections from this phenomenological analysis was used to support design of professional development sessions with teachers.

Through immersion in the data over time, an understanding of the events that occurred and what they meant to the participants in those events – including to the researcher – emerged over time. It took considerable time for the author to feel confident enough in her understanding and interpretation of the details of the interviews and events that occurred, in order to take a step back from the data in order to generalize in ways that allow portions of the events and interviews to be encapsulated as a brief story that can be told to someone who did not immerse themselves in the data and events. That challenge, of delving deeply into the data and then being able to pull back to summarize portions of the substance without the full detail, has required several years of attention and learning.

4.3. Grounded Theory

The initial approach of the study was to build theory from data in a situation where existing literature does not yet adequately provide guidance regarding how to investigate our questions. We were not able to draw on prior studies that have investigated the relationships between formative assessment practice and teachers' professional knowledge. Therefore, we designed a study that would, based on our understanding of the phenomena at hand, provide a rich data set that we could use in order to build an understanding.

In analyzing the data, we first went in with a mostly open approach. Existing literature discussed a domain of knowledge, assessment knowledge, that seemed to be a promising place to begin in understanding the knowledge that informs formative assessment. However, we believed that the domain of assessment knowledge was under-defined. Further, the nature of the domain was unclear based on what seemed to be conflicting uses of the theory in existing

literature. We therefore cast a wide net with our study design and data collection, including by giving teachers freedom to select and discuss any assessment that they thought would be useful to them, and using an open interview protocol and a semi-structured interview approach.

Although we began the data analysis with models of teacher professional knowledge from both mathematics and science in mind, our first efforts at data analysis were to do open coding of the data. Coding initially was done by hand (using a highlighter and pen, along with paper copies of the transcripts). For this initial round of coding, portions of the interviews that seemed most applicable to teachers' knowledge were coded using line by line coding as suggested by Charmaz (2006). Specifically, each statement made by the teacher was highlighted to make it distinct from other statements (we went idea by idea rather than strictly line by line on the transcript) and the key ideas were summarized in the margin. This coding was done by the author of this thesis (Millay) with support from an undergraduate student (Paige Gallagher), with both students independently coding interview transcripts from one of the assessment cycles and then discussing the codes in order to arrive at agreement regarding the meaning of the teachers' words and what was being referred to during the interviews.

During this analysis process, we found that it was challenging to confidently interpret the teachers' goals, motivations, and intentions from the interview transcripts alone. The interviews were long and at times, the teacher would make statements that seemed to be contradictory. Also, in this first case, the assessment had not gone as the teacher planned and the unfolding of events was difficult to understand through the interviews alone. Two summers and time during two academic years was spent reviewing, coding and re-coding, and discussing these initial interviews as we grappled with interpretation of the teachers' words. This coding process and the tentative codes that were being generated were shared with the thesis committee through committee meetings, and maps were generated using PowerPoint to begin

to explore relationships between the categories of statements that the teacher made. These maps were helpful in beginning to gain an understanding of the consistencies and contradictions in the interviews.

Another round of coding, with the same interview transcripts was conducted using HyperResearch. This software was used to continue the line by line coding of the same interviews and to work through mapping the data using the software. However, the problems with interpretation remained. This difficulty with interpretation led to an epistemological challenge for the researcher, leading to use of phenomenological methods, which I will describe later in this chapter. Through the use of phenomenology, it was possible to use the data more holistically in order to develop possible interpretations of the events that occurred during the assessment cycle. These interpretations were discussed at length by the author with members of the committee as well as other researchers in the community. It was very important to the author to accurately understand and represent the events that occurred in the assessment cycle. As this understanding of events developed through phenomenological study over several years, the author gained confidence in being able to trust the teachers' words as a way to capture essential elements of her thinking and gained comfort with the mix of consistencies, contradictions, conflicts, and inconsistencies that existed in the interviews and in the implementation of the assessment cycle. Through deeper understanding, it became more clear that we could learn something important about how knowledge informed this teacher's formative assessment practice, without her knowledge having to be consistent. There could be room for her to make mistakes, or to learn, or to change her mind – and we as researchers may not know the nature of the changes that occurred over the course of the assessment, because our window in to the assessment consisted of snapshots in time – the interviews and observations that we were able to conduct. This epistemological shift for the researcher – from

discomfort with the shifting nature of thought and the varied levels of understanding for which the initial coding had provided evidence – to greater confidence and comfort that these shifts represented a capturing of a true state of human existence and professionalism – allowed progress to begin again.

In the new round of line-by-line coding, the author and advisor approached coding from the perspective of existing theory. Using Magnusson, Krajcik and Borko's (1999) model of teachers' professional knowledge, the author looked for statements made by the teacher that clearly reflected elements of the knowledge defined in that model. At this stage, not all elements of the interview were coded – statements that could clearly be attributed to a knowledge domain were coded as such. Supporting statements or statements that provided insight into the quality or focus of the knowledge, were coded as well. ATLAS-ti was used to facilitate coding of Magnusson, Krajcik and Borko's paper alongside the interview transcripts from three cases of assessment with two teacher participants.

In this round of analysis, some elements of the teachers' knowledge for which there is evidence in the interviews, were ignored for the time being. For this analysis, it was sufficient to find evidence of some of the knowledge domains that informed formative assessment practice and to assemble the evidence that these knowledge domains were present. This change in analytic approach followed a recognition that in order to make progress in the research, it was necessary to simplify and narrow the initial questions. A decision was made to focus on the clearest examples of how knowledge was used to inform assessment. This decision was reasonable, given that the research community does not currently have an understanding of this topic, and that the evidence we can provide regarding some domains of knowledge makes a contribution – even if there are also other domains that play a role in assessment. These can be defined and illustrated at a later time.

4.4. Analyzing Data Using ATLAS-ti and Line by Line Coding

The author uploaded documents into ATLAS-ti, including pre- and post-instruction interview transcripts from three assessment cycles with two teachers and the Magnusson, Krajcik, and Borko article. First, she coded the article to identify key elements of each knowledge domain that was described. As examples, Knowledge of Assessment included Domains of Science Learning to Assess and Methods of Assessing Science Learning. Knowledge of Students included knowledge about common student difficulties within a domain of science learning, as well as knowledge of specific misconceptions that students might have.

After coding for the definitions of these knowledge domains (codes included Knowledge of Students and Knowledge of Instruction), the author turned to coding the interview transcripts. Statements made by the teacher that are by now quite familiar to the author were coded first according to the key ideas that were expressed, as recommended by Charmaz (2006).

The following transcript and table illustrate the coding method used. "I" is the interviewer and "T" is the interviewee.

I: So, what do you see being the subject matter or the content?

T: Well, how long, the humans have been on the earth and, well of course they would have to know basically and a general idea of when dinosaurs became extinct. Also, the general idea of when humans were here.

I: What do you see as, what is really important about the subject matter?

T: You mean the whole thing? Or just that question?

I: Well, what the question gets at. I guess…what is the question getting at?

T: Geological time, you mean? The big topic, yep, okay. So the big topic is, time. Yep, geologic time and time itself. They have a really, hard time judging how big things are and how small they are and what that really means. Like, if you asked them, 'how much room would a billion pennies take up?' It would really be hard for them, so big numbers unless you were used to working with them, kids have a hard time with them. So, just, knowing what that means.

This section of transcript was coded as follows:

Table 3. Examples of Coding Interview Transcript

Quote	Interpretation	Initial Code	Code Groups
So, what do you see being the subject matter or the content? How long, the humans have been on the earth and, well of course they would have to know basically and a general idea of when dinosaurs became extinct. Also, the general idea of when humans were here.	The teacher states that the important subject matter of the question is for students to know when dinosaurs became extinct, how long humans have been on earth, and when humans were first present on earth.	Content Students Need to Understand to Answer the Question	Knowledge of Assessment (Dimensions to Assess)

Table 3 continued

what do you see as, what is really important about the subject matter?) 10:03 You mean the whole thing? Or just that question? 10:05 (Well, what the question gets at. I guess…what is the question getting at?) 10:08 Geological time, you mean? The big topic, yep, okay. So the big topic is, time. Yep, geologic time and time itself.	The teacher states that the big topic that the assessment question is getting at is students' understanding of geologic time and the scale of time.	Big topic that the question is getting at	Knowledge of Assessment (Dimensions to Assess)
They have a really, hard time judging how big things are and how small they are and what that really means. Like, if you asked them, 'how much room would a billion pennies take up?' It would really be hard for them, so big numbers	The teacher states that students have a hard time with scale and with large numbers.	Content that is difficult for students	Knowledge of Students
unless you were used to working with them,	The teacher suggests that having students work with big numbers can help make them easier to grasp.	Not coded separately from the statement above – it is clear that the teacher is suggesting instruction (or practice) can help, but the statement is vague	
kids have a hard time with them.	But that without that work, students have a hard time with scale and large numbers.	Content that is difficult for students	Knowledge of Students

Table 3 continued

So, just, knowing what that means.	The big topic of assessment is for students to know what the large scale of geologic time means. (In this case, that means understanding that humans and dinosaurs could not have lived at the same time, given the time scales involved. This interpretation is supported by other portions of the transcript.)	Big topic that the question is getting at	Knowledge of Assessment (Dimensions to Assess)

Once coded, portions of the transcript were interpreted using the coding in order to address the research questions. In the example of transcript shown above, the teacher reflected on her choice about what domain of science learning to assess in her classroom. Her statements make clear that she has chosen to assess specific content (did humans and dinosaurs live at the same time?) that is part of a larger domain of knowledge (student understanding of geologic time, time scales, and scale in general). These represent elements of students' conceptual understanding. Magnusson, Krajcik, and Borko (1999) noted that there could be many ways of defining the dimensions of science learning that should be assessed. The example they provided described conceptual understanding, interdisciplinary themes, and nature of science as being among the dimensions of science learning that were important to assess (Magnusson et al., 1999, p.108).

In discussing this choice about the specific dimension of learning to assess, the teacher referred to her Knowledge of Students, noting that the "big topic" that was the subject of the assessment was content that she knew to be challenging for students. Specifically, she stated that students have difficulty with scale and with understanding time scales.

This method of coding and interpretation was used to systematically work through the interviews in order to inform the research questions for this study. In the following two chapters, data analysis using this method will be used to present and discuss two cases of assessment cycles. These case studies represent a focus on teachers' decisions about what to asesss and how to assess, while selecting assessment questions and while reflecting on the performance and value of those questions. During the assessment cycles that were studied, teachers made many decisions about assessment. These included choices about question/assessment selection, choices about the instruction and/or the scaffolding to provide to students as they completed the assessment, and choices about how to evaluate and respond to students' responses. At times these decisions aligned well to the goals teachers articulated for assessment; at other times, teachers changed their minds about assessment goals, or enacted assessment in ways that were inconsistent with the goals that had been articulated. Each of these choices could be analyzed separately in terms of the knowledge teachers referenced while describing the decisions they made, but doing so is beyond the scope of the current work.

CHAPTER 5

CASE 1: MISALIGNMENT BETWEEN ASSESSMENT GOALS AND QUESTION SELECTION

That question doesn't get at that piece of content – it just doesn't. (Teacher A, reflection on assessment)

5.1. Case 1 Introduction

This assessment cycle was the first assessment cycle studied with Teacher A, and was part of a unit focused on plate tectonics within the SEPUP Issues and Earth Science materials. For this assessment, Teacher A selected an analysis question from the student book. The selected analysis question was intended to be completed by students at the end of an activity about geologic time.

The procedure for the activity included providing students with a set of cards that had events in geologic time printed on them. These events included things like the formation of the Earth, existence of the first bacteria, extinction of dinosaurs, and existence of the first mammals. Students were to work in groups to sort these cards in the order in which they believed the events in geologic time had occurred. After sorting the cards, students would be provided with a reference sheet that showed the correct ordering of events.

Given the context of piloting new materials, this was the first time Teacher A had implemented this activity and the selected assessment with her students. She had, however, been a science teacher for many years and had previously taught the similar content to students at the same grade level as the classroom she was teaching at the time of the study. Teacher A also had extensive experience with assessment and had participated in more than one state-level initiative to develop assessments for classroom use. These prior experiences were discussed at length in the baseline interview.

The analysis question selected for this assessment cycle was the following:

Your younger brother tells you about a television show he watched where humans ride dinosaurs instead of cars. He says he wishes he could go back to the time when people lived with dinosaurs. Based on what you learned in this activity, what do you tell him?

The teacher's guide provided information about this analysis question, stating that the question may reveal a student misconception that humans and dinosaurs lived at the same time, and providing the following exemplar:

In real life, dinosaurs and humans did not exist on earth at the same time. Dinosaurs became extinct between 50 and 250 million years ago (about 65 million years ago). The oldest modern human fossil is only thousands of years old. That means that dinosaurs became extinct millions of years before evidence of the first modern human. Humans co-exist with dinosaurs only in fantasy or science fiction.

In the interview conducted prior to use of this assessment, Teacher A described her reasons for choosing the assessment question and her expectations regarding the information that the question would provide to inform her teaching. In the interview conducted after use of this assessment, the teacher described her observations from administration of the assessment, evaluation of students' responses to the assessment, description of follow-up completed in response to the students' responses, and reflection on the value of the assessment for informing her teaching. The teachers' reflection showed that she considered the assessment to have been unsuccessful in providing the information she had intended to elicit. The next portion of this chapter will discuss this case in light of the research questions.

5.2. Goals About What to Assess

In the pre-instruction interview, Teacher A described her goals for the assessment as being focused on gaining insight into her students' learning and about whether the instruction provided through the SEPUP materials was sufficient to (1) address a misconception that she anticipated students would have coming in and (2) support learning about a challenging topic. When asked to describe her plans for the assessment, Teacher A stated:

"What I really like about this question is they have to apply [learning from the activity] to beliefs that are wrong and are very common in the human population. Most people if you ask them that [question] a lot of adults don't know the answer to it. And what I'm curious about, why I picked it, is because I am wondering if they are going to get that, the answer to that from this activity."

In the quote above, Teacher A referred to the belief that humans and dinosaurs lived at the same time – something she anticipated many students will believe, based on watching movies and TV shows that have humans and dinosaurs together. Teacher A's interest in misconceptions and belief that misconceptions are challenging to address and to "root out" also came up in other interviews conducted throughout the study. Throughout, Teacher A indicated that she does not want to have students leave her classroom with misconceptions about science content.

While reasoning about her goals for the assessment, Teacher A explains that in the activity students will complete prior to completing the assessment, *"There isn't a card that shows a picture of humans and dinosaurs and says 'No, this doesn't happen!'"* and she goes on to state that because students are not explicitly told in the activity that humans and dinosaurs did not live at the same time, they will *"have to infer that humans and dinosaurs didn't live at the same time by doing this activity."* And she goes on to express skepticism that students will be successful in this task, indicating that her skepticism is part of her motivation for using the

assessment to check student understanding. She said, *"I really want to see if they are going to get that piece of content from this activity. I don't know that they will."*

Beyond the specific factual question of whether humans and dinosaurs lived at the same time, Teacher A saw the question as being intended to assess students' learning about geologic time, stating that, *"The problem is knowing the scope of what geological time is and they basically have to have the sequence in their mind"* in order to answer the question. During pre-instruction interview and also in the post-instruction interview, Teacher A discussed instructional strategies to support student learning about what she saw as the challenging aspects of geologic time – the scope and scale of time, the nature of time itself, and the sequencing of time.

For Teacher A, the assessment question about dinosaurs and humans was connected to the larger topic of geologic time through the phrase, "based on what you have learned in the activity", which Teacher A saw as prompting students to provide specific evidence from the activity to support their answers. When asked what evidence she expected students to provide, she noted that they could cite evidence from the geologic time cards, showing the extinction of the dinosaurs, and the time period from which the first evidence of humans was found. But Teacher A believed that to accurately use that evidence, students would need to have an understanding of how events were ordered (for example, that 250 million years ago was long before 50 million years ago) as well as the scope of those time periods (that 200 million years in between events was long enough that there would not have been any dinosaurs still around when the first humans lived).

In summary, prior to the assessment Teacher A's goals about what to assess were: (1) To assess whether students had a specific misconception after instruction that she anticipated many of them would have before instruction

(2) To assess whether the activity was sufficient to address the misconception

(3) To assess whether the activity was sufficient to provide students with an understanding of the scope and sequence of geologic time

Teacher A continued to hold these goals for the assessment in the post-instruction interview.

5.3. Plans for How to Assess

Teacher A's plans for how to assess involved a hands-off approach. Her intention was to have students complete the activity and then respond to the analysis question individually, using pencil and paper, during class time. She noted multiple times during the interview that her intention was to take a step back while students completed the assessment because she viewed the process as an "experiment" for investigating the effectiveness of the new materials and she did not want to interfere or influence her students' responses. She also stated multiple times that she "wanted to see" whether students would "get" the content from the activity – and that she was not sure that they would.

During instruction, these plans changed and Teacher A reflected on her change of plans during the post-instruction interview. She stated that as students were completing the assessment, she noticed that they were not addressing the question in the way that she had hoped they would. Students were writing stories about humans and dinosaurs, but were not explicitly addressing whether humans could have lived with dinosaurs, and were not bringing evidence from the activity into their answers. Seeing this, Teacher A began to prompt students as they were working to think about whether humans and dinosaurs lived at the same time, and to look back at the geologic time cards.

While reviewing students' responses during the post-instruction interview, Teacher A discussed the thinking that led her to intervene, in spite of her plan to stand back and observe.

She stated that the day they completed the assessment, *"quite a few kids didn't say anything about this…and I don't think it's because they didn't know it, some of them didn't know it but I think mostly it is because they didn't know it was that important. You know, somebody asks you, gives you a lecture and you are not quite sure what is important that you pick out of it. Even though we talked about it, some of the kids still didn't clue in on that, they didn't. Until I went…around and I said 'Look at the papers and make sure that they have the thing about humans and dinosaurs.' So anyone who didn't have it did have when we got done with them. So, this might not be, this isn't, it isn't exactly a true evaluation of what they learned from that lesson because if they didn't we made them do it."*

In the post-instruction interview, Teacher A reflected at length on the analysis question, stating that *"that question, doesn't get at the piece of content"* and "*I don't think it is good for the purpose for finding that out."* She went on to discuss ways that she might change the question for future use, to better emphasize both the factual piece that she wanted to assess (did humans and dinosaurs live at the same time?) and that she was looking for students to use evidence from the activity to support their answer.

5.4. Reflections on Feedback Elicited Through Assessment

In this case, Teacher A stated that her goals for the assessment were not met, and that she did not feel that it was a good assessment. She noted that she could see the effect of her intervention in some of the student responses, where students added a sentence at the end addressing whether humans and dinosaurs lived at the same time. As she went through the responses, she noted instances where it was clear that students did not have the misconception that humans and dinosaurs lived at the same time, and ones where she could not tell based on the response.

Teacher A stated, "*I didn't think they learned it yet and I was right in most cases…I still don't think they learned it."* She went on to note that to figure out whether students still believed that humans could have lived with dinosaurs, she would need to do another round of assessment.

In terms of the information that the assessment provided to support Teacher A's practice, she reflected primarily on alternative methods for assessing student learning – including follow-up questions she could ask, or ways to ask revised versions of the question that would be better at eliciting the student feedback she had looked for.

5.5. References to Professional Knowledge and Beliefs

Teacher A referenced her professional knowledge and beliefs in multiple ways as she described and reflected on her choices about what to assess and how to assess in her classroom. In this section, we will describe ways in which Teacher A referenced professional knowledge and beliefs while describing and reflecting on her selection of the analysis question she chose, using the organization provided by Magnusson, Krajcik, and Borko's (1999) model for teachers' professional knowledge.

5.5.1. Orientations Toward Science Teaching

Magnusson, Krajcik and Borko state that the domain of orientations toward science teaching "refers to teachers' knowledge and beliefs about the purposes and goals for teaching science at a particular grade level". (1999, p. 97) Their model of PCK includes orientation in a central position in relation to the other domains within PCK, showing the orientation shapes and is shaped by the other domains within PCK, as well as being informed by the three base domains (knowledge of content, pedagogy, and context.) Magnusson, Krajcik, and Borko note that

empirical studies are needed to help illustrate this theoretical role of orientations, as well as stating the prior research shows that teachers may hold multiple and competing orientations at the same time.

In analyzing transcript data, this central and shaping role of orientations became immediately apparent, as will be shown below. At 2 minutes and 19 seconds into the interview, after some initial chatting about logistics of classroom observations, the researcher asked Teacher A the following question to begin the interview:

2:19 Alright, umm so, I was wondering if we just look at the question a little bit and maybe you could tell me, sort of what you are planning to do, how you are planning to present it to your students and what will they do?

The interview proceeded as follows, with questions/comments from the interviewer in parenthesis:

2:32 Okay, umm, what I really like about this question is they have to apply it to, umm, to beliefs that are wrong and are very common in the human population. Most people if you ask them that a lot of adults don't know the answer to it. And what I'm curious about, why I picked it, is because I am wondering if they are going to get that, the answer to that from this activity.

2:57 I am not sure if they will, it is kind of an experiment for me.

3:00 (Okay)

3:00 I don't know, if they will figure it out of not.

3:03 (Okay, and what makes you question that?)

3:05 Yeah, cause it doesn't come right out and say it and kids are very literal.

This portion of transcript will be used to illustrate the role of orientation, which is also in evidence throughout the interview. The other coding applied to the quotations will also be provided to keep overall interpretation of the interview in context.

Table 4. Example of Coding Transcript for Orientation

Row	Quote	Interpretation	Initial Codes	Code Groups with Explanation
1	2:32 Okay, umm, what I really like about this question is they have to apply it	Teacher A states that she likes the question because she believes that it is a question that asks students to apply learning from the activity while answering the question (she discusses this classification of the question in depth later in the interview)	What I like about this question	Orientation –Teacher A believes there is value in having students apply learning to an assessment question. Although the instructional goal to be achieved is not made explicit at this point in the interview, the teacher's "like" of the question is guided by an internal reference to the question's intended purpose. Knowledge of Pedagogy – Teacher A classifies this as an application question. She does this more thoroughly later in the interview.
2	to, umm, to beliefs that are wrong and are very common in the human population.	Teacher A states that she likes the question because it addresses beliefs that are wrong and are very common in the human population. Later in the interview, she describes these as "misconceptions" and talks further about her thoughts on misconceptions.	Misconception definition	Orientation –Teacher A believes there is value in attending to student misconceptions. This is further supported by statements made later in the pre interview as well as in statements made in the post interview. Again, at this point in the interview the instructional goal is not yet clear – but the choice of question is clearly guided by an internal reference to the question's intended purpose. Knowledge of Students – the teacher demonstrates awareness that her students may have a specific misconception. This knowledge is aligned with the information provided in the teacher's guide provided by SEPUP.

Table 4 continued

3	Most people if you ask them that a lot of adults don't know the answer to it.	Most people, including adults, don't know whether humans and dinosaurs lived at the same time.	Misconception definition	Orientation – the prevalence of this misconception contributes to it being of interest for assessment or instruction Knowledge of Students – the misconception is common among students (and others) Knowledge of Content –the teacher does know the answer
4	And what I'm curious about, why I picked it, is because	Teacher expresses her own curiosity as a reason for picking the question	What I'm curious about	Orientation – the teacher's curiosity indicates that she is attributing value to the information this question could provide; she is seeing the assessment as a way to answer questions that she has in her role as teacher. This suggests a belief that the assessment is a way to answer questions the teacher may have, possibly about students, or about instruction. It also suggests a belief that she sees satisfying her curiosity as a way to further her students' learning. Her curiosity is a reason to pick the question.
5	I am wondering if they are going to get that, the answer to that from this activity.	Teacher expresses that she is wondering whether students will get the answer – that humans and dinosaurs didn't live at the same time – from the activity.	What I'm curious about If they are going to get that from the activity	Orientation –the teacher's curiosity implies a belief that there is value for her as the teacher in learning about what students learn Knowledge of Assessment – part of the teacher's motivation for selecting this question is that she is wondering whether students will get specific content (did humans and dinosaurs live at the same time?) from the activity. This is part of the dimension of science learning that she is choosing to assess – student acquisition of this specific science content

Table 4 continued

6	2:57 I am not sure if they will, it is kind of an experiment for me.	Teacher expresses doubt that students will get the target content from the activity and states that asking the question is a kind of experiment for her. This "experiment" refers in part to this being a new set of materials for the teacher – as she describes more fully later in the interview.	I'm not sure that they will figure that out	Knowledge of Context – the teacher refers to the context of piloting materials that are new to her Orientation – the teacher expresses doubt that students will learn the content targeted by the selected question, from the activity. It is not clear at this moment what beliefs are the foundation of this doubt, but the teacher goes on to describe some of those beliefs later in the interview.

Table 4 continued

7	3:00 I don't know, if they will figure it out or not.	Again, expressing doubt and an understanding that students should be able to figure out the target content based on doing the activity, but may not be able to do so – and that the selected question will help the teacher to see whether or not students have attained the target content.	I'm not sure that they will figure that out	Orientation – again, expressing doubt (based on beliefs not yet shared) about whether students will get the target content. Implies a belief that students will need to be actively involved in constructing meaning and will need to "figure it out" in order for the learning to be successful. Knowledge of Instruction – the statement begins to connect Teacher A's ideas about instruction with her ideas about students. In the statement, she speaks of doubt regarding whether students will be able to figure out content, based on the activity. Her knowledge regarding the substance of the activity and how the activity will further student learning within the content area, is within the domain of Knowledge of Instruction. Knowledge of Students - the statement begins to connect Teacher A's ideas about instruction with her ideas about students. In the statement, she speaks of doubt regarding whether students will be able to figure out content, based on the activity. Her knowledge regarding the ways in which students learn, preconditions for learning, and challenges to learning of the target content is within the domain of Knowledge of Students.

Table 4 continued

8	3:03 (Okay, and what makes you question that?) 3:05 Yeah, cause it doesn't come right out and say it and kids are very literal.	Here, the teacher begins to identify some of her specific beliefs about the activity and about students. The activity does not provide explicit information that will help students in answering the question. And students are literal. Therefore, students may not get the target content from the activity.	Students are literal	Orientation –the teacher expresses a belief about students (students are literal) and a belief about instruction (if students are not explicitly told something, they may not get it). This seems to shape some of her thinking within the other domains. Knowledge of Instruction – instruction may need to be more explicit than it is in the current activity (as outlined in the student book) in order for students to learn target content. This is shaped at least in part by the teacher's beliefs about the relationship between instruction and student learning. Knowledge of Students – students are literal in their understanding and may need explicit (literal) instruction in order to get target content. This is shaped at least in part by the teacher's beliefs about the requirements and process of student learning.

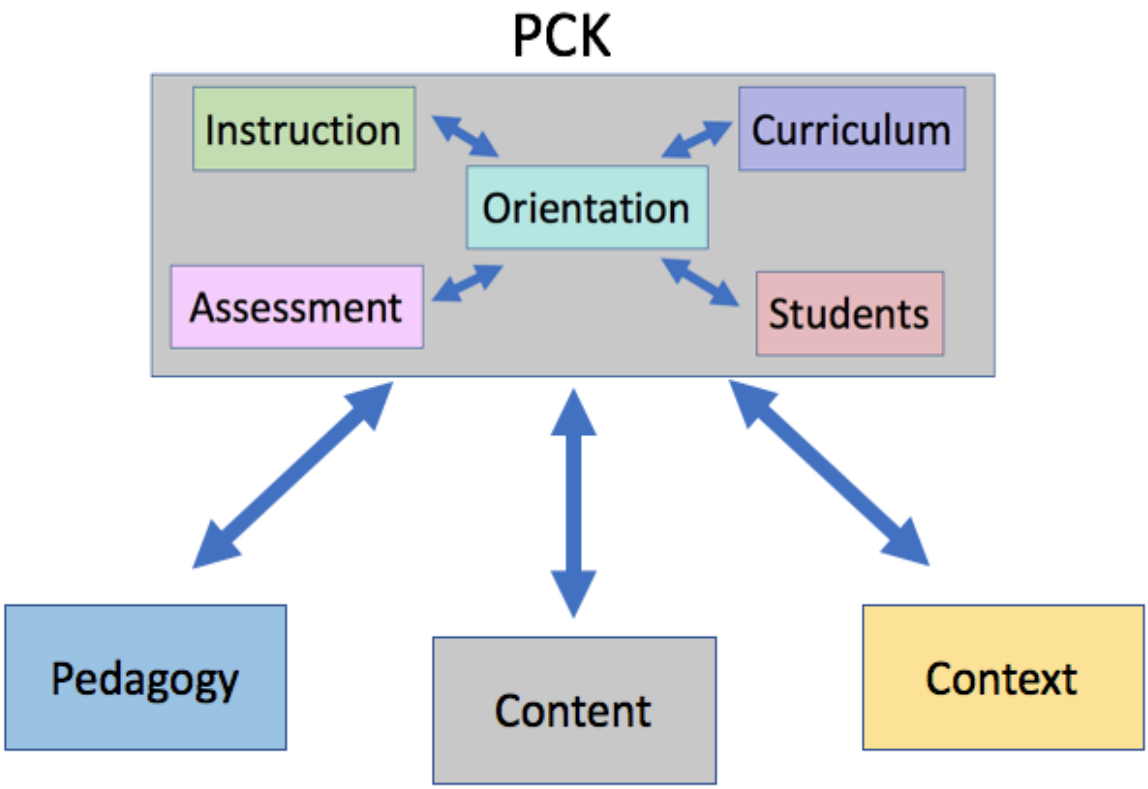
5.5.2. The Role of Orientation

As the interview continues, the teacher begins to state more explicitly what some of her beliefs are as she describes the sources of her doubts about whether students will get the content from the activity and how she may respond if they have a misconception after the activity. However, the purpose of the above illustration was to demonstrate that the presence of beliefs underlying the teacher's thinking and decisions is fundamental to making sense of the interview.

Just in the few excerpts given above, there is evidence of interplay between the teacher's beliefs about assessment and instruction, and her knowledge of pedagogy, knowledge of students, knowledge of instruction, knowledge of context, knowledge of content, and knowledge of assessment.

Magnusson, Krajcik, and Borko's model theorizes these domains and relationships, with the arrows inside PCK being described by the word "shapes" and the arrows between the base domains and PCK being described by the word "influences":

Figure 3. Magnusson, Krajcik, and Borko PCK Framework



The data above support the role that Magnusson, Krajcik, and Borko's model attributes to Orientation. For example, in rows 7 and 8, the teacher demonstrates linkage between Orientation, Knowledge of Students, and Knowledge of Instruction when she expresses doubt

that students will learn target content from the activity. The specific knowledge and beliefs are not defined in the statement, but the attribution of doubt to the likelihood of students learning target content through an activity references held knowledge/beliefs about students and held knowledge/beliefs about instruction. The teacher states that her doubt is one reason for selecting the assessment question. This suggests that Orientation is indeed playing a central role, with the teacher's beliefs (though not yet made explicit) shaping and being shaped by other knowledge domains.

There is also evidence that Orientation is influenced by Knowledge of Pedagogy, Content, and Context. For example, the teacher's knowledge of context in this case includes the knowledge and belief that she is piloting new materials in the context of seeking to evaluate them for further use. Her thinking of her assessment as an "experiment" designed to help her understand how students learn from an activity in this context, is a reflection of beliefs she holds about her role, given the larger context in which science instruction was being conducted in MainePSP classrooms. At one point she notes that: *I think it is important that we also like test the way they are because we are supposed to be piloting it.* This indicates that she is attempting to use the materials with fidelity, without inserting her own priorities, and provides an additional reason for her to use the assessment as a way to test the quality of the materials.

In another instance, value the teacher places on the selected assessment question because it asks students to "apply" their learning, is rooted in her knowledge about multiple methods of assessment and Bloom's Taxonomy (which she discusses at length later in the interview). Her knowledge of pedagogies for assessment shapes her beliefs about the purpose and value of the assessment question. From the baseline interview, it is clear that this teacher has prioritized thinking about multiple methods of assessment and has enacted this priority by

participating in assessment design efforts over a long period of time. She spoke of these efforts in detail and with enthusiasm stating, for example, *"I'm always on the pilot program of assessment because I find it very interesting."*

The baseline interview also provides some insight into the orientation behind the dimension of science learning that the teacher selected to assess in this case. In the baseline interview the following exchange illustrates that the teacher's general belief about what is important to assess about students' learning in science, tends to fall within the conceptual domain:

I: What kinds of information do you try to get from assessment?

T: Ok, for example, do they understand what I want them to understand – the important parts. Um. Do they, have they already got it and they need to do something different because they've already got it – or if they don't, what do they need to do to make their learning more complete or to improve their understanding. Ask that question again.

I: Um, what kinds of information do you try to get from assessment? What do you want them to usually understand?

T: Like the main points, the main ideas, concepts – it's not necessarily if you can memorize things to me – whether they have uh if they can use the vocabulary correctly – can they show me what they know.

Evidence from throughout the interviews supports the intertwining of beliefs and knowledge that is theorized by Magnusson, Krajic, and Borko. This intertwining makes it time consuming to separate out the role of Orientation from the other domains. Recognizing that the teacher's beliefs underlie her thinking and decision-making is helpful in interpreting the

teacher's words and decisions – it is a fundamental part of the perspective she brings to the task of teaching, to assessment, to piloting of materials, to the process of being part of a research project and of participating in an interview about assessment with a Master's student. And in general, the beliefs she brings are not made explicit, but become evident through careful consideration of the data. For example, that the teacher values providing her students with answers and guiding them to understanding, is evident in many statements throughout the interviews as well as in the teacher's actions in the classroom. It is also something that the teacher explicitly acknowledges stating, for instance: *"One of my mottos is "there's no such thing as getting it wrong". I don't let them get it wrong. I mean, I try not to."* This stance reflects on the teacher's beliefs about her role in guiding students to answers and understanding. The statement was made about the teacher's general practices of teaching during a baseline interview conducted on December 12, 2011. In another interview conducted on March 7, 2012, the teacher described guiding students toward an answer as they completed the assessment described in this case study, in spite of having resolved not to do so:

I went around and I said 'Look at the papers and make sure that they have the thing about humans and dinosaurs.' So anyone who didn't have it did have when we got done with them. So, this might not be, this isn't, it isn't exactly a true evaluation of what they learned from that lesson because if they didn't we made them do it.

Then another exchange during that same interview, while reviewing students' responses to the assessment, illustrates how the teacher's general orientation led her to guide students toward the right answer at another point, while handing the assessment papers back to her students:

I: Okay, so let's see, so, what did you do with the student responses?

T: What did I do with them? Well, I talked about them in general before I passed them out and I said 'okay, I am going to pass papers out but before I do I want to tell you I really enjoyed what you said and the important thing is, okay, answer the question before I pass this out, did humans and dinosaurs live at the same time?' and they all went 'NOOOO!!' and that wasn't the same day. So then I passed them back and then I would say to certain kids 'I love this part about the cell phones' or so. I passed them back and talked to each child if I passed it back and that's what I did with them. And I used them personally to think next year when I do that crosscheck I am going to do that differently because, either then or during the activity ***you've got to make sure that they know.*** [emphasis added]

Defining the role of orientation in shaping and being shaped by the other knowledge domains could easily be the subject of this entire thesis. However, for the sake of being able to spend time discussing the role of other knowledge domains in informing and being informed by assessment, we are going to leave this deeper study of orientation for another paper. For now, the claim made here is that orientation plays a shaping role that informs teachers' assessment practices, not only through beliefs about assessment, but also through beliefs that shape and are shaped by other knowledge domains.

The table below provides a brief summary of some of the ways in which orientation shaped and was shaped by other knowledge domains within PCK that influenced assessment decisions in this case study.

Table 5. Roles of Orientation in Shaping Knowledge

Beliefs and Intentions Shaping Assessment	Relationship Between Orientation and Knowledge Domain	Evidence
Students are likely to still have a misconception after instruction. The assessment will help identify whether students have a misconception. If the assessment shows that they do have a misconception after instruction, the teacher will work to guide them to understanding.	Orientation Shaping and Being Shaped by Knowledge of Students/Student Learning	The teacher states a suspicion that students will still have a misconception after instruction: *I suspect that quite a few kids will still believe that people live with dinosaurs.* *5:27 That's what I am thinking, that is my suspicion* The teacher states an intention to address the misconception if it is found to be present: *5:32 (And, and then what will you do?)* *5:33 I will, I will supplement in …* 5:57 *…probably, we could look on some timelines and say 'okay, people, early people didn't come until when and the dinosaurs were not here anymore, you know, when.' And let them think about it and I would guide them, I could say, I could guide them to a better understanding. But, we all know that misconceptions are hard to get rid of so it might take more than once.* Also, the teacher looks for evidence that students do *not* have the misconception while describing what she is looking for in students' responses to the assessment: *…do they have the idea that humans and dinosaurs were not here at the same time. And so, I go through and make sure that that thought is at least there. So, that's what I look for first.*

Table 5 continued

Instruction using these new materials may not be sufficient to support student learning. With other activities in these new materials, she believes that students have not learned the target content and it has been necessary for her to supplement in order to provide sufficient instruction for students – but this time she is going to run the activity as-is, and assess whether it is successful.	Orientation Shaping and Being Shaped by Knowledge of Instructional Strategies (Informed by Knowledge of Context)	*4:28 Okay, so I am wondering that there isn't a card that says picture of human and dinosaurs. No, this doesn't happen! So, they have to infer that human and dinosaurs didn't live at the same time by doing this activity and I'm not sure that they will figure that out on their own. Some of them may know it already though so that kind of, may ruin my little research experiment there. I have suspected with other activities that kids don't always learn the little piece of information or the concept of whatever it is by the activity that they are doing.* *5:08 (Uh huh)* *5:08 I am not sure of that, and I have always supplemented and done some more activities and some more explanation when they didn't get it. And so, this time I am not going to do that, I am just going to see how this comes out. I suspect that quite a few kids will still believe that people live with dinosaurs.*
There is a tension between the teacher's beliefs about how much emphasis should be placed on students' understanding of the concept of geologic time, and how much emphasis is placed on that concept in the materials being piloted. The teacher's perception of this tension has shaped her choice to assess this content, as a way to inform whether she needs to emphasize the content more than the materials do.	Orientation Shaping and Being Shaped by Knowledge of Curriculum	*Yeah, well you know to me, to me, a big concept that I'd always spend of time on is geological time and there is really not a whole lot in here [in the new materials]. So to them [the curriculum developers] maybe, I am thinking that is not a big idea but to me it is. So, sometimes it is what you believe is really important and then I am afraid they [the students] won't get it in one little dose. So to me that is important*

Table 5 continued

The teacher sees the assessment as a method to inform her about whether her students are they able to make the inference needed to answer the question. She also sees it as a method to inform her about whether the instruction guided by the activity will help students learn the content. Her skepticism that students may not learn the target content from this activity is rooted in prior experience with these materials as well as prior experience with instructing students on the same content using different materials. This skepticism guides her choice about assessment.	Orientation Shaping and Being Shaped by Knowledge of Assessment	*4:28 Okay, so I am wondering that there isn't a card that says picture of human and dinosaurs. No, this doesn't happen! So, they have to infer that human and dinosaurs didn't live at the same time by doing this activity and I'm not sure that they will figure that out on their own. Some of them may know it already though so that kind of, may ruin my little research experiment there. I have suspected with other activities that kids don't always learn the little piece of information or the concept of whatever it is by the activity that they are doing.* *5:08 (Uh huh)* *5:08 I am not sure of that, and I have always supplemented and done some more activities and some more explanation when they didn't get it. And so, this time I am not going to do that, I am just going to see how this comes out. I suspect that quite a few kids will still believe that people live with dinosaurs.* Later in the interview the teacher notes that in her prior teaching, she spent considerable time on the concept of geologic time. This prior experience and comparison of past practice to what is in the materials to be piloted, leads to a belief that students may not learn the target content from the activity and is one basis for her assessment choice: *Yeah, well you know to me, to me, a big concept that I'd always spend of time on is geological time and there is really not a whole lot in here [in the new materials]. So to them [the curriculum developers] maybe, I am thinking that is not a big idea but to me it is. So, sometimes it is what you believe is really important and then I am afraid they [the students] won't get it in one little dose. So to me that is important*

In the examples above, the blurring of lines across the domains within and beyond PCK is evident. For example, when speaking of tension between her own goals [Orientation] and the likely goals guiding the design of the instructional materials being piloted [Knowledge of Curriculum], the teacher's meaning references multiple domains. These include her choice to have students complete an assessment of a specific domain of science learning [Knowledge of Assessment] that will help her determine whether instruction was sufficient [Knowledge of Instruction] to overcome what she believes will be a prevalent misconception in her classroom [Knowledge of Students].

5.5.3. Knowledge of Pedagogy

Teacher A's Knowledge of Pedagogy is referenced in Line 1 of the first table in this chapter, when she notes that to answer the question she has selected, students need to "apply" learning to content about which they may have a misconception. She states that, *"what I really like about this question is they have to apply it to…beliefs that are wrong and are very common in the human population."* This suggests that her choice of question in this case was influenced by her knowledge of pedagogy for assessing students. In the baseline interview, Teacher A talked extensively about strategies and methods for assessing student learning. Later in the pre interview, she discusses how to evaluate the level of thinking that an assessment is asking students to use by looking at the kinds of words used to prompt them to respond. She takes out a sheet that refers to elements of Bloom's Taxonomy along with verbs that prompt students to provide answers drawing on different levels of thinking, and begins to walk through the levels:

T: Okay so, lower order skills are here and you notice we work our way up. So there are, there are the words you would ask kids to do. These are all recall level, for kids to do. And the next one is, 'how did this happen?, Explain this.' That's the comprehension level.

T: Third one, you are applying it to something, construct something, draw it, give me an example, illustrate. Now, analyze your, you know, you're taking it apart. And then here, evaluating, you say' okay, given these two things, given these two choices, which one would you chose and why?' Perhaps, or, yep, 'what is your opinion? And you have to support it.' There are those words again.

Later in the interview, the teacher provides evidence that this pedagogical knowledge about strategies for assessing student learning informs her general assessment practices, stating:

T: *You look for verbs but remember they could be tricky because sometimes they are not.*

I: Is this something? Do you use this?[referring to the sheet the teacher had taken out that had levels of thinking associated with verbs used to ask questions of students]

T: Yes, I do.

I: How do you use this?

T: When I am doing, planning a unit, I try to make sure I have some of everything in there

Teacher A goes on to review the selected question in light of the information on the sheet. Her understanding of the level of this question is that it asks students to "apply" what they have learned from the activity. At another point she uses the word "infer". She also speaks at length about the need for students to provide evidence from the activity in their answer to the question. For example, in speaking about what she will look for in evaluating student answers, she states that:

T: I'll be looking for, evidence from okay, talk about the activity. And I might even, if they don't talk about in this activity, they I am going have to say, 'Okay, tell me in this activity.' Even if they get it right, 'tell me in this activity how you would of figured that out.' Cause that is what the question is…

While discussing what she expects students to do with the prompt, Teacher A notes that the type of question she has selected in this case is familiar to her from questions she has selected in other contexts, stating that:

I have done this kind of problem a lot. 'What would you, like, cousin Sandy thinks that you know that, rocks all come from one place or whatever. How would you convince cousin that they are wrong?'

And that as part of the prompt, she might remind her students that they also are familiar with the type of prompt, because she has used these kinds of prompts in the past:

So that's the kind of thing I would say. 'Remember we had this kind of a prompt before and how you handle it. What would you say?'

The teacher's Knowledge of Pedagogy, including knowledge of methods of assessment and experience with the type of prompt used in this question, contributed to her decision about method of assessment in this case. She notes that in reviewing student responses, she will be looking for evidence that students have applied learning from the activity in their responses to the question, stating: *I am going to look at their answers and I am going to see if they do indeed come out with, some kind of evidence that shows that they know that dinosaurs and people didn't live at the same time.*

In the post interview, when the teacher reflects on the assessment, her negative evaluation of the assessment focuses back on this issue of whether the selected question was

appropriate. In the interview, she described walking around the room as students were responding to the assessment question and noticing that students were focused on the story-telling element of the question and not on the question of whether humans and dinosaurs could have lived at the same time and the evidence they might provide to support their answer. She then described prompting students to address that aspect of the question by directing them back to the activity and her belief that by prompting the students, she led them to the answer, therefore undermining the value of the assessment as a tool for evaluating students' learning from the activity. She noted examples of student responses where students added that humans and dinosaurs didn't live at the same time, on to the end, stating that students had added that information because she had told them to. And she discussed the challenge of knowing whether students didn't provide information that got at the core of the question because they didn't know the information (or had a misconception) – or because they didn't know that it was important to provide that information in response to this prompt.

"I just think it was just not a good question, because it was specific enough 'based on what you learned in this activity', I or the translator or the book, whatever needs to be really specific about when organisms lived instead of leaving it wide open like that."

In this case, the teacher's pedagogical knowledge from prior experience informed her choice of assessment, in that she selected a prompt that seemed familiar to her and that she expected would assess whether students could apply their learning. However, in this case she found that the prompt did not serve her intended purpose and that to get the types of responses she wanted from students, either the written prompt or her verbal instructions needed to be more specific about the kind of answer and evidence that students were expected to provide. In this case, the failure of the prompt to serve the teacher's purpose turned out to be crucial, because it meant that the teacher was not able to answer the questions that she had

intended to answer with the assessment, about whether students had a specific misconception and whether the activity was sufficient in supporting student learning about a topic that she considered to be important.

About the assessment, she stated:

T: I wouldn't say it was a very good evaluation.

I: So you wouldn't say it was a good one?

T: If you have to tell them what the answer is, generally it is not good. So, I don't know, that question, doesn't get at the piece of content, it just doesn't. Unless, you came right out and say to them like, 'Based on what you learned in this activity about geological time, based on, if you learned about, based on what you learned about when human beings came into existence or when dinosaurs became extinct…' so you might have to be more specific.

From the teacher's perspective, poor quality in the prompt meant that the question didn't "get at the piece of content" that she was trying to assess. In this case, the teacher's knowledge of assessment pedagogy from prior contexts informed her choice about how to assess students in a new context. But, her knowledge from prior contexts turned out not to be sufficient enough to ensure success in using the prompt for her purposes in this context. Her reflection after the assessment was therefore focused on her Knowledge of Assessment, noting that in this case, different pedagogy in the form of some modification to the prompt would be needed in order to elicit the responses she wanted from her students.

These findings about the role of pedagogical knowledge support the relationship theorized by Magnusson, Krajcik, and Borko, with Knowledge of Pedagogy informing and being informed by PCK – in this case, informing Knowledge of Assessment. The process of doing the

assessment then led the teacher to expand her Knowledge of Assessment as she gained knowledge about how to assess student learning in this new context. The process also led the teacher to reflect on her pedagogy, noting that: *"You don't always get at what they are thinking, you don't. Unless you ask, you got to pull the right things out, you got to get the right things to get what you want. I have always, I am pretty good questioner and I am listening for what they are saying and I am pretty good a probing and getting out what I need to have come out so that would need a little more, a little more of that probing part."* The lack of success of the assessment pedagogy in this case was brought up by the teacher again in the subsequent assessment cycle, in which she compared the success of the two cycles, noting that the second was more successful than the first. Given the teacher's extensive experience with assessment, it is not clear whether the teacher actually increased her knowledge of pedagogy through the assessment cycle, or whether the process simply led her to reflect on things she already knew.

5.5.4. Knowledge of Students

In this case, Teacher A's knowledge of students shaped her choice of assessment question, and was shaped by her beliefs about students. She believed that her students were likely to have a specific misconception (that humans and dinosaurs lived at the same time) and that they were likely to have difficulties understanding the scope of geologic time. As a result, she selected an assessment question that would help her identify whether students had the misconception that she expected them to have, and whether they were able to use evidence from the activity (such as the length of time between extinction of dinosaurs and evidence of the first humans). Based on students' responses, she intended to deliver additional instruction to support student learning.

Teacher A's Knowledge of Students was referenced in Line 2 of the first table in this chapter, when the teacher states that: *"what I really like about this question is they [students] have to apply it [learning from the activity] to…beliefs that are wrong and are very common in the human population."* In this statement, the teacher defines what she sees as misconceptions, as "beliefs that are wrong and are very common in the human population". In the statement above, the teacher begins to articulate her expectation that her students will have a misconception that humans and dinosaurs lived at the same time. This theme of misconceptions is one that comes up several times in both the pre- and post-instruction interviews for this assessment cycle. Teacher A sees addressing misconceptions as part of her role.

Later in the pre-instruction interview, Teacher A describes some of her reasoning for how students may have come to believe that humans and dinosaurs lived at the same time. For example, she states that:

"There are so many movies, popular culture that these kids watch that have dinosaurs and people together."

Knowledge of common misconceptions is one area of Knowledge of Students as Science Learners, according to Magnusson, Krajcik, and Borko. In this case, whether the belief that humans and dinosaurs lived to together truly is a misconception depends on the definition applied to misconceptions. In some instances, a misconception is simply defined as a commonly held incorrect idea. The SEPUP materials being piloted by the Teacher A through the MainePSP assert that the belief that humans and dinosaurs lived at the same time, is a misconception. Other definitions of misconceptions may require that the incorrect idea be connected to a specific conception that is incorrect – for example, a formed conception about geologic time that specifically allows for humans and dinosaurs to have lived at the same time, where the conception about geologic time would be the misconception, with the belief that humans and

dinosaurs lived at the same time as an implication of the larger misconception. But regardless of whether the teacher's definition in this case would have matched up with MKB's definition of misconceptions, the teacher clearly expected students to have a specific difficulty with the content; a difficulty that the SEPUP materials also anticipated students would have.

Teacher A also noted another common difficulty that she anticipated her students would have with the content, stating that: *"[Students] have a really hard time judging how big things are and how small they are and what that really means."* At another point, in talking about scale, she stated that, *"their math conceptions are, they are not really up to par"* in reference to students' challenges with scale. In this and other statements, Teacher A described what she anticipated students' challenges with understanding the scope of geologic time would be. This is another area of MKB's Knowledge of Students – specifically, the knowledge of common difficulties that students may have with the content. This challenge of students understanding the scope of geologic time is documented in existing literature.

Teacher A's expectation that students would be challenged with understanding the scope of geologic time, together with her expectation that some would come to her classroom with a misconception that humans and dinosaurs lived at the same time, contributed to her selection of the assessment question. Going back to her first statement from the pre interview, Teacher A stated that what she liked about the assessment question was that to answer the question, students had to apply learning from the activity to an area where they might have a misconception. Later in the interview, the teacher described other likely challenges for students, in the form of challenges with understanding the scope and scale of geologic time. She saw these challenges as making it likely that students would have difficulty in answering the assessment question.

Her understanding of what students would have to do in order to answer the assessment question correctly was that, *"They have to infer that humans and dinosaurs didn't live at the same time by doing this activity and I'm not sure that they will figure that out."* The teacher describes her curiosity about how students will answer, whether they will demonstrate that they have a misconception, or whether they will be able to use the activity to arrive at a correct answer, as a reason for her selection of the assessment. As additional evidence that finding out about her students' thinking was a reason for question selection, the teacher talks at other points in the interview about how she will respond to her students' answers. At one point she notes that prior to administering the assessment she doesn't "know what their misconception is going to be yet" and that she is waiting to see what it will be, in order to tailor further instruction.

5.5.5. Knowledge of Instruction

Teacher A described her knowledge of instruction as a key element in shaping her choice of what domain and topic of science learning to assess in this instance. Specifically, she notes that there is a tension between how the materials being piloted address the topic of geologic time, and how she typically has addressed it in the past. She states that: *"A big concept that I'd always spend a lot of time on is geological time and there is really not a whole lot in [these new materials]."*

Teacher A goes on to describe that in general, in her experience with the materials being piloted, she has believed that the content was not addressed effectively enough to support student learning. She stated that: *"I have suspected with other activities that kids don't always learn the…concept …by the activity that they are doing."*

These two elements of the teachers' knowledge of instruction played a role in the teacher selecting an assessment question that was focused on the topic of geologic time; (1) that geologic time is a topic that requires time to teach and (2) that the piloted materials do not contain a significant amount of material focused on geologic time and, in general, sometimes do not sufficiently address target content. These two pieces of knowledge about instruction interacted with the teachers' knowledge about students – knowledge that students struggle with the concept of geologic time. Together, the teachers' knowledge of instruction and of students in this situation led the teacher to believe that students were likely to have difficulty in learning about geologic time and that they would require more instruction than was given in the materials being piloted. The teacher went on to reflect about ways in which she might approach providing additional instruction, depending on what she learned from the students' responses to the assessment question.

In thinking about how she would support student learning after the activity, she stated that: *"We could do, there are a few different things. I have exactly thought what I am going to do because I know basically what I do, but I don't what their misconception is going to be yet. But probably, we work on some timelines and say 'okay, people. early people didn't come until when and then the dinosaurs we're not here anymore and you when.' And let them think about it and I would guide them, I could say, I could guide them to a better understanding. But, we all know that misconceptions are hard to get rid of so it might take more than once. I will have to tailor it to that.*"

This statement reaffirms that the teacher anticipates that students will have difficulties in answering the assessment question and that she intends to address those difficulties through additional instruction. The instruction that she considers to be most likely is to provide additional support to students in using a timeline to consider whether it would be possible for

humans and dinosaurs to live at the same time. Also in this statement, her expectation that the assessment question will provide her with information that she will be able to use to inform her instructional response, is clear. She states twice that she will tailor the follow-up to students' misconceptions, depending on what those are after instruction.

The teachers' expectation that students will have difficulties, her belief that the concept of geologic time is important for students to grasp, and her concern that the instruction provided in alignment with the new materials will be insufficient, shaped her choice of assessment question in this case. In discussing why she selected the question that she did, she stated that: *"I really want to see if they are going to get that piece of content from this activity. I don't know that they will."* Her statement shows that she expects the assessment question to reveal what students know about the content, and that she anticipates that students will have difficulty in understanding the target content based on the activity.

5.5.6. Knowledge of Curriculum

In this case, there is some evidence of both tension and alignment between the teachers' goals for instruction, and the goals enacted through the materials. The teacher's perception of tension contributed to her choice of assessment question. In describing her reasons for selecting the question, she notes that she personally places more importance on the topic of geologic time, than is placed on the topic in the current materials. She states: *"Yeah, well you know to me, to me, a big concept that I'd always spend of time on is geological time and there is really not a whole lot in here. So to them maybe, I am thinking that is not a big idea but to me it is. So, sometimes it is what you believe is really important and then I am afraid they won't get in one little dose. So to me that is important."*

In her statement, she specifically notes that at times, her beliefs as a teacher about what is important for students to learn can be at odds with the focus of the curriculum developers. This tension contributes to the teachers' belief that the instruction to be provided using the materials faithfully, may not be sufficient to bring about student learning. Not only does she see less emphasis in the materials ("not a whole lot in here") than she would have included in her prior instruction; but she sees this as a matter of differences in priorities, between her belief that the topic is important for students, and what appears to be the belief of the curriculum developers, that it is not as important. Because Teacher A believes that the emphasis on the topic is less than her emphasis would be, and that the instruction to be provided is less than she would normally provide, she states that she is skeptical that students will learn the content she intends for them to learn. She also takes a stance toward the assessment as "an experiment" to see whether the instruction provided in the materials is sufficient for student learning. Her skepticism and wondering about whether students will learn the content from the activity provided, in the absence of other instruction, motivated her to select a question that would probe student learning of this content.

At the very start of the interview, the teacher notes that, *"what I'm curious about, why I picked it, is because I am wondering if they are going to get that, the answer to that from this activity. I am not sure if they will, it is kind of an experiment for me."* Again, this statement combines her thinking about the curriculum and the priorities of the new curriculum (in tension with her own), together with her knowledge of instruction (that significant instruction is needed to help students learn this topic) together with her knowledge about students, (that they are likely to have difficulties). These concerns motivate her choice to assess this topic using this question.

In addition, there is alignment between the teachers' perception of goals and the goals of the curriculum, in encouraging students to use evidence. The teacher discusses evidence at length during the interview as a key part of what students will need to do in order to answer the question. She also notes that use of evidence is a goal of this curriculum, and of curricula being used in her school, stating that: "*all the time they are being asked to, all the time, 'give evidence about what you think and why.' Cause then it really forces you to go back and think about what you know. Instead of just saying 'yes' or 'no' if you have to back it up. Now, that something we stress at this school, in there writing they have to explain and back it up with evidence. Everything that they do, they have to back it up with evidence in math, in you know, everything. You can't just say 'you're wrong'.*"

So her knowledge of curriculum supports her selection of this assessment questions in two ways. (1) It provides her with an opportunity to test whether she is correct, that the lesser emphasis of the materials on the topic of geologic time will result in a lack of student learning that she will need to correct with additional instruction and (2) it asks students to engage in an activity that is a priority both for the science curriculum, and for the curricula used in the school in other subject areas. Both of these avenues reinforce that the material of the question is important to assess.

5.5.7. Knowledge of Context

In this case, the teacher was cognizant of having a responsibility to pilot materials with fidelity, given her role as part of an improvement community. About the materials, she stated that, "*I think it is important that we also test the way they are because we are supposed to be piloting it."* This situation of piloting materials as part of an improvement community provided important context to the teacher's choice of assessment question.

Specifically, the teachers' knowledge of context (participating as part of a pilot and the purpose of the pilot) combined with her beliefs about that context (that it is important for her to implement materials with fidelity in spite of her suspicion that materials might come up short in some respects) contributes to the teachers' stance toward the assessment as an experiment and as a situation where she intends to implement the materials and assessment as they are, rather than supplementing, until she learns about what her students have learned from the activity. She states that her intention is that: *"I will try to keep my mouth shut too and not talk too much about it and not say 'this is what they mean' and I really want to see if they are going to get that piece of content from this activity. I don't know that they will."*

Again in this statement, her skepticism that the instruction guided by these new materials will lead to student learning is evident and she states this as part of her motivation for selecting the assessment question, stating that *"I really want to see if they are going to get that piece of content from this activity."*

5.5.8. Knowledge of Content

In this case, knowledge of content played a lesser role in Teacher A's decision about assessment. Primarily, knowledge of content informed the teachers' thinking about students and instruction. While in discussing her choice of assessment question Teacher A, immediately brought up her thinking about students (that they would have a misconception). She also immediately brought up her thinking about the attributes of the question (that students would need to apply learning in order to respond); and her thinking about the instruction students would receive (that it might not be sufficient for teaching the target content). But she did not discuss the content that the question was asking about until prompted to do so by the interviewer.

When Teacher A did begin to discuss the content of the question, it was first in the context of how she expected her students to answer the question, in response to the interviewer asking about what kind of evidence from the activity students might be able to use to support their responses. She responded as follows: "Well, how long, the humans have been on the earth and, well of course they would have to know basically and a general idea of when dinosaurs became extinct. Also, the general idea of when humans were here. Because that's, I often don't agree with having kids memorize you like, exactly 3 million years ago and there's even discussion among scientists when you can say that humans were here. But no one, that I know of, no one says that dinosaurs and humans were here at the same time. Nobody, of that I know of!"

In the later part of this statement, Teacher A states her belief that whether humans and dinosaurs were alive at the same time is not controversial in the scientific community. She also states that her goal is for students to understand the general idea of when dinosaurs became extinct and when humans were first alive, and that she does not see the exact time periods to be as important for them to learn as the general idea.

The interviewer then asked the teacher for more information about the content behind the assessment question. In the exchange below, the teacher explains that she considers the key content to be the nature of geologic time, and students' abilities to understand how vast it is. She also sees students' use of evidence to be a key part of the content behind the question.

I: "Well, what do you see as, what is really important about the subject matter?"

T: "You mean the whole thing? Or just that question?"

I: Well, the question gets at. I guess, what is the question getting at?

T: Geological time, you mean? The big topic, yep, okay. So the big topic is, time. Yep, geologic time and time itself. They have a really, hard time judging how big things are and how small they

are and what that really means. Like, if you asked them, 'how much room would a billion pennies take up?' It would really be hard for them, so big numbers unless you were use to working with them, kids have a hard time with them. So, just, knowing what that means. Gee, I just thought of something else too. Of course, sequencing in those kinds of ordering of date.

I: And then you mentioned also, the idea of evidence and how they would figure out from the activity.

T: Right, yep, so that is important too that, they really answer the question because it asks them 'what you learned in this activity?' , not just what you know. So, if somebody know it and they say, they leave that part out, I will have to say "Okay, could you tell me, could you write that.'

The questions asked in this interview were not sufficient to probe the extent of the teachers' content knowledge about these topics. But from the way the teacher discussed content in the interview, a general knowledge of the content – that humans and dinosaurs did not live at the same time, that geologic time is vast, and about what constitutes evidence that students could bring forward to address the questions – were a backdrop to the teachers' thinking about the assessment question. However, her selection of this assessment question in order to assess the content was more a matter of her curiosity about her students and her uncertainty that the instruction would adequately deliver the key content, than it was about the content itself.

Further, given the nature of the question and the challenges that arose when the teacher administered the question and tried to evaluate students' responses, it turned out that the question did not meet the teachers' goals in terms of having students address the core content, which the teacher described as the scope of geologic time, the evidence that humans and dinosaurs did not overlap in geologic time, and effective use of evidence from the activity. So in this case, the teachers' knowledge and beliefs about students, instruction, and pedagogy

led her to select a question that turned out to not be effective in assessing content. This may have been a result of the teacher now having enough in-depth knowledge of the content to allow her to recognize that the question would not prompt for the right content knowledge. But given that the teacher demonstrated having the content knowledge that she was seeking in her students, it is more likely that the breakdown occurred in one of the other knowledge domains.

5.5.9. Knowledge of Assessment

Knowledge of assessment, according to MKB, is knowledge of the domains of science learning to assess, and methods for assessing those domains, within a specific content area. In this case, the teachers' knowledge of assessment was demonstrated by the choices she made about assessment. The outcomes of the assessment – in this case, an unsuccessful outcome from the teachers' perspective – shows that he knowledge of assessment in this case was not fully formed. The knowledge she had at the outset was not sufficient to bring about a successful outcome. This is not surprising, given that it was the first time this teacher was using this assessment question, and that it was the first time she was teaching using the pilot materials. One might therefore expect that she would be missing the content and context-specific knowledge about what to assess and how to assess it in this case.

In selecting the domain of science learning to assess in this case, the teacher selected three specific elements of the content domain and one element of the cognitive. She intended for the question she selected to assess the overarching concept of geologic time, and stated that, "*The problem is knowing the scope of what geological time is and they basically have to have the sequence in their mind"* in order to answer the assessment question she had selected. The way she intended to determine whether students had the correct sequence and an understanding of scope, was through assessing a specific scientific fact – whether humans and

dinosaurs lived at the same time. The teacher notes that this fact may be one about which students have a misconception – therefore putting a burden on the activity to show students that their misconception is incorrect. For example, she states: *"What I really like about this question is they have to apply [learning from the activity] to beliefs that are wrong and are very common in the human population."* In the prior statement, the teacher notes that students have to "apply" learning while answering the question. She later states that the nature of the activity and of the assessment will mean that they have to "infer" that humans and dinosaurs didn't live at the same time: *"There isn't a card that shows a picture of humans and dinosaurs and says 'No, this doesn't happen!' - so, [students] have to infer that humans and dinosaurs didn't live at the same time by doing this activity."* She considers these cognitive activities of applying and inferring, to be important to assess. Her method for assessing whether students have accomplished these cognitive activities is to establish whether students are able to take evidence from the activity and use it to support their answers. She states that, "this asks for specific evidence based on what you learned in this activity. So, there will be other evidence or they might just say 'I know' or 'I saw a movie' or 'I read a book' or 'I already learned this', that is evidence too but that will be separate from this. I want them to be able to tell me in this activity and I will look for that." Use of evidence – a part of one of the NGSS practices – is therefore also a part of the content domain that the teacher decided to assess, as well as being part of the method for assessing whether students gained understanding of the other elements of the content domain.

The method Teacher A selected for assessing the domains she had selected to assess, was an open response question in a format that was familiar to her from prior use. She had previously demonstrated knowledge of many different options for assessing student learning – including using concept maps, exit slips, quizzes, and other strategies. In this case, she selected

an assessment question that she anticipated would require students to demonstrate knowledge of the domain she intended to assess.

However, in this case, the assessment did not go as planned. Prior to completing the assessment, Teacher A noted that: *"What I really like about this question is they have to apply it [what they learned in the activity] to beliefs that are wrong and are very common in the human population."* She went on to describe that the question asked students to use evidence from the activity to support their answer and that they would need to have an understanding of the sequence of events in geologic time, and some understanding of the scope of geologic time, in order to provide a complete and accurate answer. But after the assessment was administered, the teacher re-evaluated the question she had selected and concluded: *"I just think it was just not a good question, because it was not specific enough…that question, doesn't get at the piece of content"*

Her reasoning in determining that the question was poorly worded and did not get at the domain she had intended to assess, included her observation that during administration of the assessment, she noticed that many students were responding to the question with stories. But that they were not addressing the core content that she had intended for them to address. She responded to that situation by prompting students to address the content, reminding them that the question was about whether humans and dinosaurs lived at the same time. For example she stated that during instruction, *"I said to the kids who still didn't get it I said 'when you had those cards out and what happened when, were dinosaurs and humans ever there at the same time?' and they go 'no', 'okay, so what does that have to do with this question?' So it was like relating what was in the activity to that answer."*

In evaluating the students' responses, she noted that several students had simply added a note to the end of their response, stating that humans and dinosaurs never lived at the same

time. The teacher was dis-satisfied with these responses, because the students did not bring evidence from the activity in the way that she had expected. She also felt that she was unable to determine whether students had the misconception that the question was designed to probe, stating "some of them you can't really tell". And as a result, she was uncertain about the effectiveness of the activity in bringing about student learning. Rather than relying on the assessment question to provide her with feedback in order to augment instruction, she ended up relying on her general feeling about the quality of the activity, noting that she believed that the students who knew the answer going in still knew it; and the ones that did not, still did not. "*I didn't think they learned it yet and I was right in most cases. I had to, I still don't think they learned it. The kids that knew it already, still know it...*"

How can we interpret this outcome? In this case, the teacher did not have specific assessment knowledge that involved use of this assessment question to assess the domain she was seeking to assess. However, she had relevant knowledge from other contexts that she applied in selecting the question. In this case, she combined her knowledge about the domains of science learning that she wanted to assess with knowledge about the new context. Specifically, knowledge of the goals of the curriculum, the possible gaps in the instructional materials, the challenges students might have with the content, and the kinds of assessment questions that might be productive in assessing student learning. Given that the teacher had not used this question in the past, she relied on her knowledge from prior assessments in order to evaluate the quality of the question. Based on her understanding of Bloom's Taxonomy, she believed that the question would assess at the cognitive level that she was looking for. Based on prior use of refutation questions, she anticipated that the question would lead to students' using evidence to support their answers. These assumptions based on prior use of other assessment questions turned out to not be true in this case and the teacher concluded that in a

future year, she would use a different question – perhaps one she wrote herself – or would re-word the question to better meet her goals.

5.6. Summary

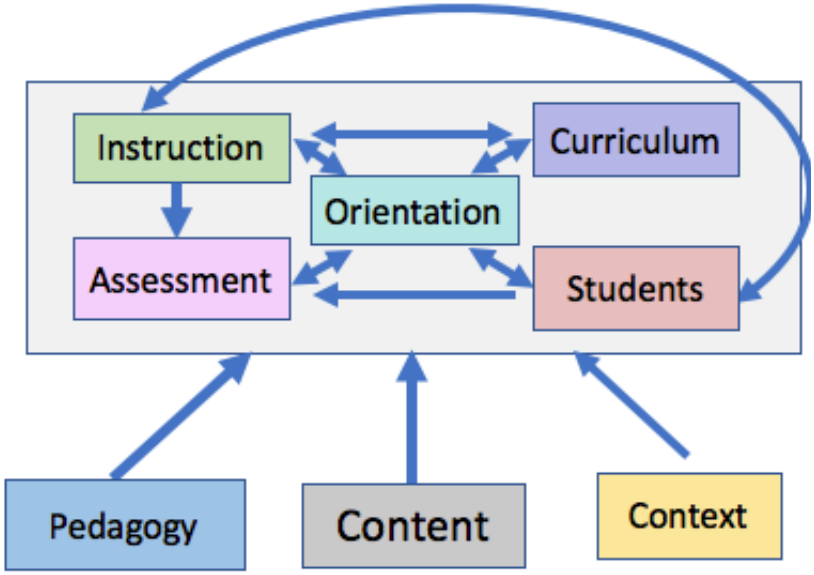
In this case, Teacher A referenced knowledge from across domains while describing and reflecting on her thinking about goals for what to assess and plans for how to assess student learning. For example, she referred to knowledge of students (likely misconceptions, likely difficulties) and knowledge of instruction (kinds of instruction needed to teach the content) while discussing her belief that students might have difficulties with the content domain she selected as the target of assessment. Teacher A described being motivated to assess the areas she intended to assess based on her perception of their instructional importance, her belief that students would have difficulties with those areas, and her skepticism that the content was adequately addressed in the new materials. The context (piloting new materials) and a tension the teacher perceived between her goals for instruction and the goals evident in the materials being piloted, contributed to Teacher A's experimental stance toward the assessment.

When selecting a method for assessing the chosen domain, Teacher A drew on prior experiences with using similarly worded questions. But in this case, Teacher A found that the question was not aligned with her goals. While reflecting on students' responses, she noted that students did not address the question in the way that she had intended, and that their responses therefore did not provide her with the information she had been looking for. She considered options for remedying this in the future, including re-wording the question to be more explicit in asking that students address the factual question she had wanted them to address, as well as being more explicit in asking for evidence from the activity.

Teacher A had intended to gain insight into whether students had a specific misconception about humans and dinosaurs that persisted after instruction and, therefore, into whether the activity was sufficient to support student learning about this aspect of geologic time. But, students did not respond to the question in a way that she believed provided her with that insight. Rather than being able to gain an understanding of whether the instruction provided was sufficient to bring about student learning of the target content, the teacher found herself reflecting on the assessment itself and on how to improve the assessment question for future use.

Why did this unexpected outcome happen, in spite of the considerable thought and reflection that went into planning this assessment and the teacher's considerable length of experience with instruction and assessment? One lens for looking at this outcome is through Magnusson, Krajcik, and Borko's (1999) model. By separating out the elements of knowledge that Teacher A referenced while describing and reflecting on her goals and plans, we can see the complex interplay of these elements. Knowledge and beliefs about students and instruction, an experimental stance toward the piloting of new materials, a perceived tension between the teacher's goals for students and the goals of the new materials, as well as personal curiosity, were all referred to by Teacher A as contributing to her goals for what to assess. In terms of the domains referred to while considering plans for how to assess students, Teacher A primarily referenced her knowledge of what types of questions to use for assessing student learning. A summary of the knowledge elements that Teacher A referenced and how some of those were referenced, is provided in Figure 4. Arrows in the figure represent elements informing or influencing each other.

Figure 4. Elements of Knowledge Referred to by Teacher A



Teacher A was drawn to the question she chose because it assessed for an interesting misconception and because the wording of the question was similar to others she had used in the past, in that it used a scenario designed to elicit refutation of an incorrect idea. Based on her initial plan to stand back and have students complete the assessment without her input so that she could gain insight into the effectiveness of the materials, Teacher A left the initial work of scaffolding student responses entirely to the selected question. This approach turned out not to meet her goals – so much so that she stepped in as students completed the assessment to guide students in responding to the question in a way that better aligned with her goals, in spite of her explicit plan to stand back. In spite of stepping in and guiding students (and in part because of how and when she stepped in) Teacher A felt that the assessment ended up not being a good evaluation of what students had learned in the activity. Students initially did not know what aspect of the question was important to respond to; when prompted, some provided the information that Teacher A wanted, but without providing a level of detail that would show her that they understood the thinking that would lead to a correct response. The

end result was a misalignment between Teacher A's goals for the assessment, and the methods she used for assessing student learning. And the mechanism of the misalignment appeared to be that the teacher did not anticipate the difficulty students would have in responding to the question in the way that she wanted. Upon reflection after the assessment was complete, the teacher considered options for re-wording the question to make it more clear. She also considered additional pedagogical steps she could take to elicit student thinking either verbally or in writing. But during the assessment cycle, both the question itself (as Teacher A suggested) and the scaffolding provided to students (in the form of prompting them toward the type of answer she was looking for), fell short of the teachers' goals for eliciting student responses. So in spite of investing considerable thought and bringing considerable knowledge to bear on the enactment of assessment, a misalignment of goals and methods, possibly caused by a breakdown in the method of assessment (the question) and the pedagogy used to support assessment (the scaffolding) led Teacher A to find this cycle to be unsuccessful.

An implication of this is that aligning goals and methods of assessment may play an important role in the assessment process. Without alignment between goals and methods, the goals may not be met – and the consequence may be that a teacher is left without the information about her students and about instruction that she had hoped to gain from the assessment. This represents a sacrifice of time and effort for teacher and students. So, what knowledge supports a teacher in evaluating their own plans for assessment and checking to see that goals and methods are aligned? Case 2 provides a contrasting example of alignment between goals and methods.

CHAPTER 6

CASE 2: ALIGNED ASSESSMENT

6.1. Introduction

In this assessment cycle, Teacher B selected an analysis question from an activity in the SEPUP curriculum focused on the history of the theory of continental drift. This activity was part of the same unit as the activity in Case 1; a unit focused on plate tectonics. Like Teacher A, Teacher B was an experienced science teacher who had taught middle school science for over 10 years. Teacher B had participated in professional development during the summer prior to piloting the new materials. As part of that professional development, she had worked through the materials that would be taught during the pilot and had developed a document showing the alignment of the materials to the Maine state standards. Where she and colleagues identified gaps, they had added suggestions in to the SEPUP materials to help them meet state standards.

For this assessment cycle, Teacher B initially selected the following analysis question:

Imagine that you have been asked to write an encyclopedia entry about the movement of the earth's continents. Write a paragraph about continental movement, describing the history of this idea and citing as many pieces of evidence as you can.

During the pre-instruction interview, Teacher B discussed wanting most of her students to respond to the question, starting with a blank piece of paper. In the SEPUP materials, a writing frame was provided as an option to use in scaffolding the question for students. The writing frame prompted students to provide elements of the history – who came up with the idea of continental movement, when the idea was developed, and what the evidence was to support the idea. Teacher B initially planned to provide the writing frame as an option for students who might be challenged with the writing portion of the assessment. However, during

the post interview, the teacher noted that she had given it as an option to all of her students, and that most had elected to complete the writing frame. However, she did not see this change as being crucial to her assessment purpose. Her summary of the assessment cycle was that it was a good check-point for seeing whether students understood the theory of continental drift and could provide evidence taken from prior activities to support the theory. Overall, she felt that her students had done well with the assessment and that it was helpful in preparing them, and in helping her to prepare them, for the end of unit assessment.

6.2. Goals About What to Assess

During the pre-instruction interview, Teacher B characterized the assessment she had selected as, "…a stopping point to…see what they know about continental drift. Because if I haven't stopped there, I don't want to get to plate tectonics and have them mixing and melding and having misconceptions that they got from this point. So, it was just a really easy place to say yes let's stop and look here, what do they know about continental drift, what have they got, and what don't they have. She went on to state that, on "our final assessment, one of the pieces they need to be able to do is explain continental drift, and the evidence that support continental drift. And so, this is an obvious stopping point before we start into plate tectonics…" These statements indicate that her goals for the formative assessment she had selected, were connected to her goals for assessment of the entire unit – and that the alignment of the goals of the formative assessment to the end-of-unit assessment, was one of her considerations in selecting the assessment.

In terms of what Teacher B planned to look for in students' responses, she stated that, "I'm going to look first to see…what they say about continental movement and then continental drift…get that theory first. And then, I'm going to see what they can give for evidence." She

also stated that she wanted to see that students knew who first came up with the idea that the continents had moved over time. And she elaborated that students might use pieces of evidence regarding the fossil record that they learned about during the two lessons prior to the assessment. She noted that the materials were new to her and that she was curious to see what evidence the students would cite, and that she would be mentally comparing their responses to responses gathered using a similar assessment during prior years of instruction with other instructional materials. Like Teacher A, Teacher B intended to gain insight into student learning, as well as into the quality of the instruction using new approaches recommended by the new materials.

In summary, prior to the assessment Teacher B's goals about what to assess were:

(1) To assess whether students could state the theory of continental drift

(2) To assess whether students could state the name of the person who developed the theory

(3) To assess whether students could cite evidence that supported the theory, based on the activities that they had completed in class

(4) To assess whether instruction using the new materials was as effective on this topic as the instruction she had done in the past, and whether it was effective enough in helping students to understand the evidence supporting the theory, or whether she would need to supplement this understanding prior to the unit test

Teacher B continued to hold these goals for the assessment in the post-instruction interview.

6.3. Plans for How to Assess

As mentioned previously in this chapter, Teacher B initially intended to provide the analysis question to students without scaffolding. She noted that a writing frame was made available by the publisher of the materials, but she was interested in using the more open-ended version of the question because she believed that by the middle of the school year, her students should be prepared to respond to a question of that type. She was initially curious about how many pieces of evidence students would provide, if asked to provide as many as they could. But, she considered the merits of the writing frame, because the writing frame asked students to respond directly to the specific items that she wanted them to respond to (what the idea was, who came up with it, and what the evidence was for the idea). Also, the writing frame asked specifically for three pieces of evidence – which was the number of pieces of evidence that she had stated she would tell students to provide, if they asked her how many pieces of evidence to provide in response to the question – and she anticipated that they would ask.

Ultimately, Teacher B provided the writing frame as an option for all of her students, and most used the frame rather than free-writing a response. She believed that either way – the open-response question or the writing frame – would meet her goals for the assessment, but felt that the writing frame was probably a more efficient way to get at the information that she was looking for from her students.

During the pre-interview, Teacher B discussed some aspects of learning that she wanted to assess that, she noted, the question she had selected would not assess for. She then went on to consider additional ways in which she might assess elements of student learning that would not be captured with this assessment method.

In reflecting on her methods after students completed the assessment, Teacher B was satisfied that the method she used was successful in helping her meet her assessment goals.

6.4. Reflections on Feedback Elicited Through Assessment

Teacher B expressed satisfaction with the responses students gave to the assessment. She noted that the assessment had provided the check on student understanding that she had wanted at this point in instruction, and that it had provided her with opportunities to provide feedback to students who had difficulties with the assessment. The most prominent difficulty she noted was that some students used a single piece of evidence more than once, rather that citing three separate pieces of evidence. But overall, she was satisfied that students had gained an understanding of the evidence that supported the idea of continental drift through the prior activities, and that they were able to pull those pieces of evidence from the materials they had worked with to date. Based on the assessment, she did not see a need to re-teach the material and did not discuss a need to revise instruction on the topic of continental drift for future years.

6.5. References to Professional Knowledge and Beliefs

Teacher B referenced her professional knowledge and beliefs in multiple ways as she described and reflected on her choices about what to assess and how to assess in her classroom. In this section, we will describe ways in which Teacher B referenced professional knowledge and beliefs while describing and reflecting on her selection of the analysis question she chose, using the organization provided by Magnusson, Krajcik, and Borko's (1999) model for teachers' professional knowledge, as we did in the previous chapter with Case 1.

6.5.1 Orientation

In this case, Teacher B focused the assessment activity on an area of student learning that would be assessed summatively at the end of the unit. *"…our final assessment, one of the pieces they need to be able to do is explain continental drift, and the evidence that support*

continental drift. And so, this is an obvious stopping point before we start into plate tectonics…This is a stopping point to say I'm going to check in to see what they know about continental drift. Because if I haven't stopped there, I don't want to get to plate tectonics and have them mixing and melding and having misconceptions that they got from this point. So, it was just a really easy place to say yes let's stop, look here, what do they know about continental drift, what do they got, and what don't they have."

Throughout the pre-instruction interview, she discussed the need to reinforce content with students and to provide them with opportunities to learn how to successfully respond to the kinds of questions they would be asked on the end of unit assessment. This approach is aligned with her focus on standards-aligned instruction, which she described in-depth in both the baseline interview and the pre-interview. In her view of standards-aligned instruction, it is the job of the teacher to look at standards and at the content students will be held accountable to at the end of a unit of instruction, and to review plans for instruction to ensure that students are given opportunities to learn and practice those things that they will be held accountable to. If gaps are found either in advance or during instruction, the teacher needs to find ways to bridge those gaps, so that students have an opportunity to succeed. Teacher B views learning as a developmental process, in which the teacher needs to ensure that students are being given goal-appropriate and age-appropriate instruction, and that learning of important concepts is reinforced through exposure, practice, and feedback. Students are then held accountable to learning targets and standards. She describes the process of going from standards to the end of unit assessment, as follows: "*We looked from our standard first…and said, O.K., it's talking about the lithosphere and the cycle interactions within that, and how earth has changed over time, is that evolution/fossils piece. And so we said, O.K., if that's our standard, what are the, what are the big ideas in that standard. And so then we, said O.K. well the big ideas are continental drift*

and plate tectonics, and our convergent and divergent boundaries, um, you know, our different types of boundaries. And so, sort of, we took it from the standard level first. And then, broke it, the standard down, and then unpacked the standard and said, O.K., once we unpack it, what does it look like."

During the interview the teacher discussed providing age-appropriate targets for students, in the context of describing a time when she assessed her students at a level that was inappropriately difficult for them. She stated that at that time, *"my poor 6th graders were assessed on things that they never should have been assessed on at 6th grade level".* At another point in the interview, she noted that what was appropriate for her to assess was related to the time of year and to what students would have learned up to that point in the year. She stated: *"So, at this point in the stage of the game, they're in February, I want to see, O.K., give me your constructed response without giving, without me giving you a frame."* In this statement, she refers to her expectation that since students have been given similar assessment tasks and have had opportunities to practice providing evidence and constructing responses to questions earlier in the year, they should be able to do so without scaffolding in this instance.

In the interview segment below, Teacher B describes how the formative assessment selected in this case relates to the end of unit assessment, and what students will be asked to do on that assessment. She goes on to state that the assessment she has selected in this case is "an obvious stopping point" before moving on to additional content, and refers to the assessment as an opportunity to "check in" with the students and "*give them [the students] feedback because it really is formative, it's not for a score".*

She states that the content she plans to assess is *"the building blocks"* for student success on the assessment and that the purpose of the formative assessment *"is to find out if*

there's places that we just need to review or go over. Or, if they're all set it, and for some kids I may say O.K., you know this is on your test and you can't give me the evidence, so here's, let's make some quick flash cards, continental drift, what is it. Here's your evidence. So, they have it ahead of time to study."

She also describes a clear alignment between the content she selected to assess, and the overall goals of instruction that she is moving her students toward. In response to the following question: "how do you see this sort of, set of concepts fitting within the, the whole unit . . . as far as, yeah, what are the other important concepts that they will be learning in the unit?" she states matter-of-factly: *Oh, it's the building blocks, I means because you have continental drift there, and then you have the plate tectonics there which then goes into the earthquakes and volcanoes…So, it's that whole piece that sets the stage for saying wait a second . . . the earth always didn't look like it does now. And so it's a very basic building block for them to get that concept and understand that."*

Teacher B's views about science and scientific knowledge parallel her views about student learning, in that she sees the process of building scientific knowledge as a developmental one. She considers it important that her students gain an understanding of how scientific knowledge has built over time and through the contributions of multiple individuals, and how ideas that seem to fail, can contribute to idea development over time. *"I fully believe in the whole idea that you pass on knowledge, from generation to generation, and you share that piece, and sometimes, I just, I think that's important for them to see sometimes. I may feel like I failed in this or that aspect, but it might be my grandchild that picks it up and says oh wait this wasn't a failure, this is just how she had to tweak it or fix it… sometimes you don't get to see any reality from your stuff until, you don't get to see it. Someone else picks it up and runs with it . . . and makes it go further…*

These over-arching approaches toward standards-aligned, developmental instruction focused on building student understanding of content toward an end-goal of student competence, made the assessment selected by the teacher *"an obvious stopping point"* for checking student knowledge and reinforcing ideas. The teacher described the process of selecting the assessment from the available options as a process of looking for *"what would work for me",* and stated that she selected this assessment rather than developing one of her own because the question got at what she wanted to assess and that she assumed, because it was part of the materials, that the question had been previously vetted with students, as follows:

I: Um, so this one question I had is, a, why did you choose to use a SEPUP question rather than designing your own for this particular assessment?

T: Because it was there and it looked like it was going hit the same thing I wanted it to hit. So, there was no point for me to create my own in this instance if it's already there. And obviously they've tested… they've used it on other kids before, so you, I would assume that the wording was clear for kids.

A difference between this case and the one discussed in Chapter 5 is that Teacher B articulated a coherent, consistent strategy for determining what to assess in her classroom. Her approach in selecting the assessment question used in this case was to start with standards and expectations for student learning, and to unpack those and walk backward to determine what students needed to learn and practice. She describes the assessment question she selected as being what she would have chosen to assess, had she been writing her own assessment. The way the teacher described the process of selecting an assessment question was as a process from standards and unit assessment, to instruction, to formative assessment to support instruction – and that in looking through the student book, she found "what would work" for

her, given her goals. In Chapter 5, Teacher A articulated selecting the question in that case based on curiosity about what students would do with the question. She later connected the question to a larger topic and set of goals; but the initial focus in selecting a question to use for assessment was on the question itself, and the specific features of the question, rather than on those larger goals. So although both teachers connected the formative assessment used in these cases to the larger goals of instruction, the priority of the specific question, versus the larger goals, was more on the larger goals in this case, and more on the specific question in Case 1.

6.5.2. Knowledge of Curriculum

Knowledge of curriculum played a prominent role in the interviews with Teacher B and in how she described her approach to assessment, both in general and in this case. As described above, Teacher B discussed her process of reviewing, unpacking, and aligning assessment and instruction to standards as an overarching approach to instruction and assessment. In order to accomplish this unpacking and alignment, she worked over the summer to align the new curricular materials to standards and worked additionally with a partnering teacher at her local school to develop summative assessments that were standards-aligned and to identify and fill any gaps in instruction that she found in the new materials, prior to teaching the material. She stated: "*We looked from our standard first…and said, O.K., it's talking about the lithosphere and the cycle interactions within that, and how earth has changed over time, is that evolution/fossils piece. And so we said, O.K., if that's our standard, what are the, what are the big ideas in that standard. And so then we, said O.K. well the big ideas are continental drift and plate tectonics, and our convergent and divergent boundaries, um, you know, our different types of boundaries. And so, sort of, we took it from the standard level first. And then, broke it, the*

standard down, and then unpacked the standard and said, O.K., once we unpack it, what does it look like."

She went on to note that this process, *"takes a lot of practice. That takes a lot of practice going through…because you first you have to know what's behind the standard and you only get that from reading the stuff of how the standards are built. And then you have to say O.K. now I know that's behind the standards, this is sort of, if I'm taking it piece by piece…"*

This discussion of how to unpack standards and how to align assessment to standards and align instruction to assessment, was a theme that came up in each interview with Teacher B. In addition, the researcher was invited to participate in an assessment design session during which Teacher B worked with her partnering teacher to develop an end-of-unit assessment. During that work session, the two teachers reviewed the standards, discussed them in depth, wrote assessment questions, and worked their way through the student materials that would be used for instruction. As they completed that process, they identified specific portions of standards that would be assessed, discussed specific wording of assessment questions in order to prompt students to address the standard, and looked for opportunities during instruction for students to gain knowledge and practice the skills that they would be tested on. Where gaps were identified, the teachers discussed specific ways to supplement instruction in order to provide students with the knowledge they would need to use while answering the assessment questions.

Teacher B also described having "internalized" the process of unpacking standards as a result of having done it for more than 10 years. She also described how standards had changed over time, referring to the Framework that proceeded NGSS (which were not yet released at the time data were gathered for this interview).

The process of aligning standards, assessment, and instruction played a key role in guiding the assessment that Teacher B chose in this case. In describing why she selected this assessment, she noted that it gave students an opportunity to practice something that they would be assessed on a the end of the unit. *"Well, one of, when we, our final assessment, I don't have it here with me, but our final assessment, one of the pieces they need to be able to do is explain continental drift, and the evidence that support continental drift."* She also stated that the question at the end of the activity on continental drift was *"an obvious stopping point before we start into plate tectonics…This is a stopping point to say I'm going to check in to see what they know about continental drift. Because if I haven't stopped there, I don't want to get to plate tectonics and have them mixing and melding and having misconceptions that they got from this point. So, it was just a really easy place to say yes let's stop, look here, what do they know about continental drift, what do they got, and what don't they have."* She goes on to state that **"**it was a smart, it felt like a good stopping point", reflecting that the selected assessment aligns well with her plans and goals for instruction.

6.5.3. Knowledge of Content

In this case, knowledge of content also played a much more prominent role than in the case in Chapter 5. While discussing the standards and the end-of-unit assessment, the teacher also discussed the content at length, in terms of specifically how the content that she would assess with the selected question, was related to the content that she wanted students to learn. Her discussion of the content was coherent and demonstrated that her selection of this assessment question, rested on her knowledge of content and her goals for instruction. Specifically, the area of content that she intended to assess was the theory of continental drift – what the theory was, what the evidence for it was, who the idea came from, and (as a bonus)

why the idea was not accepted by scientists at the time that it was raised. She saw these elements of the content as being important to assess, because after the activity on continental drift, students would move on to an activity about the theory of plate tectonics, which would provide an underlying explanation for the evidence that Alfred Wegener used to develop the theory of continental drift. Her overarching goal for instruction as of the assessment in this case was for students to understand the evidence supporting the theory of continental drift and also, to understand the history of the idea. She saw this as a crucial building block that would help them to understand the theory of plate tectonics while allowing them to see the evolution of ideas in science, which she saw as a crucial part of the content. Teacher B's confidence with the content was such that she and her co-teacher had added a full lesson to the unit, to teach students about seafloor spreading because she believed that they needed to have an understanding of that process in order to understand what drives tectonic motion.

This deep knowledge of the content contributed to the teacher believing that the assessment was *"an obvious stopping point before we start into plate tectonics…This is a stopping point to say I'm going to check in to see what they know about continental drift. Because if I haven't stopped there, I don't want to get to plate tectonics and have them mixing and melding and having misconceptions that they got from this point. So, it was just a really easy place to say yes let's stop, look here, what do they know about continental drift, what do they got, and what don't they have."*

She noted that this content was important to assess because: "*it's the building blocks, I mean because you have continental drift there, and then you have the plate tectonics there which then goes into the earthquakes and volcanoes…So, it's that whole piece that sets the stage*

for saying wait a second . . . the earth always didn't look like it does now. And so it's a very basic building block for them to get that concept and understand that."

And she described how both the processes of scientific discovery in the natural world, and the social processes of theory development over time, contributed to current understanding of plate tectonics – and that these ideas were ones that were important for students to understand as well. For example, in talking about the lesson she added to the new materials about seafloor spreading, she noted that: *"…we're going to have a day in here where we just talk about the seafloor spreading that was discovered when they looked at the different molten rocks in the seafloor in the 1950's and 60's because we think that's a piece that fits before plate tectonics. And so, we want to make sure that we, and I know they touch on it here, but we to make sure that we give the kids that information so they have that information as well, so that they see it. It's a whole continuous process that theories are built on other theories before they come to fruition. So, it's that whole, building idea of science."*

Teacher B also described her understanding of part of the nature of scientific knowledge and her desire for students to understand how knowledge develops in the scientific community. In the passage below, she describes wanting her students to understand the history of the theory of continental drift, and how an idea that was initially rejected, contributed to development of an accepted theory.

"I would like them to be able to tell me why people said no, no, no Wagner, his idea, didn't fly, and this is why it didn't fly. I'd like them to be able to tell me that. Um, I don't know if they'll have, get that or not. But I would like them, I would love it if they knew that. Because sort of that's, that's the sort of piece that . . . leads into, O.K., his idea didn't happen in science, you know, sometimes you don't get to see . . . any reality from your stuff until, you don't get to

see it. Someone else picks it up and runs with it . . . and makes it go further…because I fully believe in the whole idea that you pass on knowledge, from generation to generation, and you share that piece, and sometimes, I just, I think that's important for them to see sometimes I may feel like I failed in this or that aspect, but it might be my grandchild that picks it up and says oh wait this wasn't a failure, this is just how she had to tweak it or fix it…

Teacher B drew on her extensive knowledge of the content, including both the scientific ideas and how the ideas came about, together with her focus on the learning she wanted to foster in her students, in making her decision about what to assess in this case.

6.5.4. Knowledge of Students

Teacher B's knowledge of students played a contributing role in her decision about what assessment to use with her students. Throughout the pre interview, Teacher B referred to her expectations about what her students should have learned, based on their grade level and point in the school year. One example was when she discussed the pros and cons of using a writing frame to scaffold the assessment. She noted that she was "on the fence" about using the writing frame and described her reasoning:

I: It sounded like you kind of have a trade-off in mind about, you were thinking about doing the writing frame, but you decided not to. What things were you considering?

T: Well, because kids . . . sixth graders always do better when they have, sort of, they have the bulleted, like this, is outlined for them. And, it gives them their framework to, in which to write. And, so I was thinking, oh that would be really good, because then they could have, it would be really quick, you know, this is the idea, this is, it would be very easy scoring because I could obviously see if it's there. And so that was, those were the pros of using it. And then the

con was, well, what if they have something they want to tell me that doesn't actually fit in here. Or, what if, because they have all of this, it leaves them too close to, draws, they say oh I need to have evidence, here's my three pieces of evidence. Whereas if I said here you need to write this out. I want to, want to see how they can form their answer because I have given them a lot of a lot of frames.

T: So, at this point in the stage of the game, they're in February, I want to see, O.K., give me your constructed response without giving, without me giving you a frame. And, they had done it . . . um, when they did their boomtown choice, they had done it without a frame. And they did really well.

In the passage above, Teacher B describes her expectation that at this point in the year, students should be able to construct a response that uses evidence and reasoning without needing the scaffolding of a writing frame – but she notes that students at this level will do better on the assessment if given that structure. She also notes that some students will still *"struggle if I don't give them a starting point."*

Teacher B also demonstrates that she believes that students need reinforcement and practice in order to learn. At several points, she discusses her intention to have the assessment serve as an opportunity for her to assess whether her students understand the theory of continental drift and whether they can identify evidence to support the theory. If they can't, she intends to provide additional instruction and support, either to the entire class (if many students demonstrate difficulties) or to the individual students (if only some of the students demonstrate difficulties). "*If I see patterns, we'll talk about it as a class. If I see, with, it's just one or two kids, I'll put, usually what I'll write, and what I'll write on there is I'll put on the side make sure you come see me, or see me at study hall, or we need to talk at study hall."* About the assessment,

she stated that her goal was: *"to find out if there's places that we just need to review or go over. Or, if they're all set it, and for some kids I may say O.K., you know this is on your test and you can't give me the evidence, so here's, let's make some quick flash cards, continental drift, what is it. Here's your evidence. So, they have it ahead of time to study."* This statement reflects her intention to use the assessment formatively to support reinforcement of ideas that students will be assessed on later. It also reflects that the assessment was selected in this case to inform her about how her students were thinking about the content at this point in instruction so that she could address difficulties with them before building toward other learning.

Evidence from the post interview shows that Teacher B was satisfied with her students' performance on the assessment and saw it as useful for informing her teaching and their learning. She described students as having done well with the assessment and noted that most of the class was able to identify the theory, who had the idea for the theory, and evidence supporting the theory. Enough of the class was successful that she concluded she did not need to re-teach or reinforce for the whole class, and instead gave individual feedback to students who had difficulties.

6.5.5. Knowledge of Instruction

Teacher B demonstrated considerable confidence about her instructional strategies for teaching the content that was the subject of the assessment. Although the materials were new to Teacher B, she had previously taught the same content with students. She noted that *"we've always…talked about continental drift and Alfred Wegener. And, it's always been, like we started with a puzzle, like we always did, sort of like this puzzle was, but then the next day I would tell a story about Alfred Wagner. And it really would be like a story time, and the kids always got such*

an amazing understanding of it. And they, they could tell us all the things because they made all these personal connections.

Teacher B then went on to describe a stance of testing out the new materials, that was similar to that described by Teacher A in the previous case, noting that: "*And so when I saw this after doing these two activities, activity 40 and 41, I was like I want to know how much they get by just going through these two procedures."* Later in the interview, Teacher B affirmed that in part, she was checking to see whether the instruction using the new materials was as effective as her prior instruction in supporting student learning.

"Because in the past we've always said, O.K. now we're going to figure out what you know about continental drift, tell me what you can, retell the story to me, and they're always very good at retelling the story. And so, I'm just curious, I, I sort of want to see how, how it goes. What will have to be added…I just want to see that piece…because right now I'm really stepping back and saying, O.K., I'm not going to stress, make sure you look at that evidence, because I thought, O.K., well this has the evidence piece, when they go through and they do their checkmark and they analyze the evidence, they'll have those pieces. So, this is really them talking in their group, and it's very much not a teacher directed, it's not a lecture at all. It's really group conversation, here's our stuff, let's figure it out and go from there. So, I'm curious to see."

In the passage above, Teacher B described how in prior instruction, she would assess to see whether students could tell the story of the history of the idea of continental drift, and that with past instruction, they were successful in doing that. She noted that the new materials approach the instruction differently than she had in the past, and required some adjustment on her part, stating that she was "stepping back" because she recognized that the new materials were student centered. And she stated that she was "curious to see" whether students would

be able to identify the evidence and tell the story, based on the group discussion and the material provided in the current activities.

However, later in the interview, Teacher B agreed that the assessment that she selected from the student book to use in this case was similar to what she would have had students do in the past, when working with other materials, with only some differences in wording.

I: Um, so this one question I had is, why did you choose to use a SEPUP question rather than designing your own for this particular assessment?

T: Because it was there and it looked like it was going hit the same thing I wanted it to hit. So, there was no point for me to create my own in this instance if it's already there. And obviously they've tested. They've used it on other kids before, so you, I would assume that the wording was clear for kids.

I: Yep. So and it was, sounds like it's a little different from what you've done before?

T: Yeah.

I:. …as far as some of the wording mostly.

T: Yep, yeah.

I: But, it's basically the same task.

T: Right.

Teacher B noted that her goal for the assessment was to determine whether students needed additional support in order to prepare for the end of unit assessment, stating that her goal: *"is to find out if there's places that we just need to review or go over. Or, if they're all set it, and for some kids I may say O.K., you know this is on your test and you can't give me the*

evidence, so here's, let's make some quick flash cards, continental drift, what is it. Here's your evidence. So, they have it ahead of time to study."

She also noted that if she found that students struggled with the assessment, she would reinforce the ideas with additional instruction, both in the current year and in a future year, stating: *"…for next year when I do it, what are the things that, if I get all their entries back, and I read them, and I say oh none of them picked up on this evidence, then I know that when we do that activity I'm going to write a note in my book to say stress this piece to make sure that that happens. So, for next year when I teach it, and even for this, to go back and say, O.K., let's look at this again real quickly, so, they have that piece."*

Teacher B's knowledge of instructional strategies contributed to her decision about what to assess in this instance, because she had previously assessed similar learning while teaching students about the same topic. The instruction to be conducted in this case was different in some respects from the instruction Teacher B had done in the past, and one purpose of assessment was to investigate whether the instruction using the new materials would be effective and sufficient in getting students to the understanding that Teacher B expected of them. If not, Teacher B intended to follow up with re-teaching and feedback for students as needed, and to modify instruction as needed for future years.

As noted above, Teacher B considered this to be a successful assessment and expressed satisfaction with her students' learning through the activities in the new materials. She noted that they took evidence from the two prior activities to contribute to their responses and that most of her students were able to be successful with the assessment. For those who had difficulties, she provided them with individual feedback – but she did not identify significant shortcomings in the instruction that had been provided.

6.5.6. Knowledge of Context

Teacher B's consideration of the context of piloting new materials was discussed above under Knowledge of Instructional Strategies. Specifically, Teacher B noted that the material being taught was the same topic with the same goals as what she had taught previously, but that the approach was different. The difference she noted was that it was student-centered, relying on a group of students to work through an activity and identify evidence. In terms of the assessment, this led Teacher B to wonder whether the current instruction would be as effective as her prior instruction in supporting the same student learning that was typically her goal.

Unlike Teacher A (who had primarily cited shortcomings in student learning in prior units and activities) Teacher B noted that during an earlier unit, students learned effectively through the new approach in the piloting materials. She articulated a hope that they would do so in this unit as well, stating: *"I'm hoping, like, it will be just like the rock cycle where they came back and they were like oh we know this, this, and this. I'm hoping that happens. I'm just curious to see if it happens."* Like Teacher A, Teacher B articulated a curiosity about whether she would find evidence of adequate student learning using the instructional approaches supported by the new materials.

6.5.7. Knowledge of Pedagogy

Teacher B's general knowledge of pedagogy was woven in to her Knowledge of Instructional Strategies and Knowledge of Students in this case. Specifically, Teacher B had an understanding that students learn through a developmental process and that providing clarity about expectations along with reinforcement and feedback as students complete tasks, is

effective for bringing about student learning. Her work to align instruction and assessment to standards in this case also reflected an overall pedagogical approach of standards alignment, that was supported by the context of her school and school culture, as well as by her work with the other science teacher in her building.

Like Teacher A, in the context of using new materials Teacher B was drawn to an assessment question that was similar to assessment approaches she had used in previous instruction. In addition, Teacher B considered the pros and cons of using a writing frame to help scaffold student responses to the assessment question. Initially, she favored not using it, because she expected that students should be able to complete the task without it. One reason she noted for not using the writing frame was to allow students to add things from their own thinking that might not be prompted for by the scaffolding. *"[The writing frame] gives them their framework to, in which to write. And, so I was thinking, oh that would be really good, because then they could have, it would be really quick, you know, this is the idea, this is, it would be very easy scoring because I could obviously see if it's there. And so that was, those were the pros of using it. And then the con was, well, what if they have something they want to tell me that doesn't actually fit in here."*

Ultimately, she provided the scaffolding as an option for all students, and most elected to use the scaffolding. Teacher B considered that use of the scaffolding led to a more efficient way for her to assess what she wanted to assess, ultimately prioritizing the content of what she wanted to assess, over the process of students developing the structure for themselves, or being able to be more creative than the writing frame allowed.

While discussing her goals for the assessment, Teacher B initially noted that one of her goals was for students to recognize how the theory of continental drift led to the current theory

of plate tectonics. As she described this goal, she looked closely at the prompt that she was providing to students, and noted that the prompt did not actually ask students to take that step in their thinking. This recognition demonstrated her knowledge of the relationship between the wording of the prompt and the student responses that she could expect. As a result of noting that the prompt did not take student thinking as far as she ultimately wanted for it to go, she suggested a future assessment that she might do with students to take that next step. Below is an excerpt from the conversation where Teacher B recognizes that one element of what she wants students to do, is not actually captured in the assessment prompt that she intends to use:

"I think the nature of it lends itself to more of a retelling of mastery type stuff of, um, having this is what I've learned, this is what I just learned, this is what I'm telling back to you. I think it just lends itself to that. I would be curious to see if any of them get to the point where they are going to say and this is how . . . And, the prompt doesn't lead it to that."

"You know, that's the part of the conversation that I would love to have them say O.K. And, so how do you think this contributed to where we are now. And maybe that's the piece that it comes back to. We say, I say pull this back out now, now that we've done plate tectonics, now that we've done this, how has continental drift led to plate tectonics."

In the excerpt above, the teacher considers her pedagogy for assessment, including the details of the prompt that she will use with students. As with other elements of her instruction and assessment, she checks the alignment of the prompt with her goals for students – and notes that there is a lack of alignment. She then goes on to consider ways to meet her goal using an additional assessment, again providing evidence of her overall strategy of using instruction and assessment to build student understanding in alignment with her end goals. In this case, her knowledge of pedagogy leads her to tailor her expectations for the assessment to the result that

she expects from students, given the prompt she plans to use. She considers that her additional expectations that do not fit with the prompt, are ones to address with an additional prompt at another time.

6.5.8. Knowledge of Assessment

Teacher B elected in this case to assess student learning of specific science content that was aligned with the content she planned to assess on a summative, end-of-unit assessment that she had previously developed in alignment with both standards and the planned instruction. Although she discussed a goal of supporting student learning about how scientific knowledge develops (nature of science), the assessment she selected did not specifically prompt for that thinking, especially when students used the writing frame. The primary focus of Teacher B's assessment was on whether students could correctly state the theory of continental drift, whether they knew who had the idea that became the theory, and whether they could correctly state evidence that supported the theory.

The method Teacher B selected for assessing this domain was an analysis question from the student book. Although Teacher B initially considered providing most students with a blank piece of paper and asking them to address the question, she ultimately offered all students the option to use a writing frame that specifically prompted for the theory, the person who came up with the theory, and the evidence behind the theory. She recognized that this represented a trade-off in terms of student creativity.

The method Teacher B selected for assessing student learning was one that was familiar to her from prior instruction on the same topic, with different materials. While discussing the assessment, Teacher B noted that the selected prompt did not meet her goals for assessing the

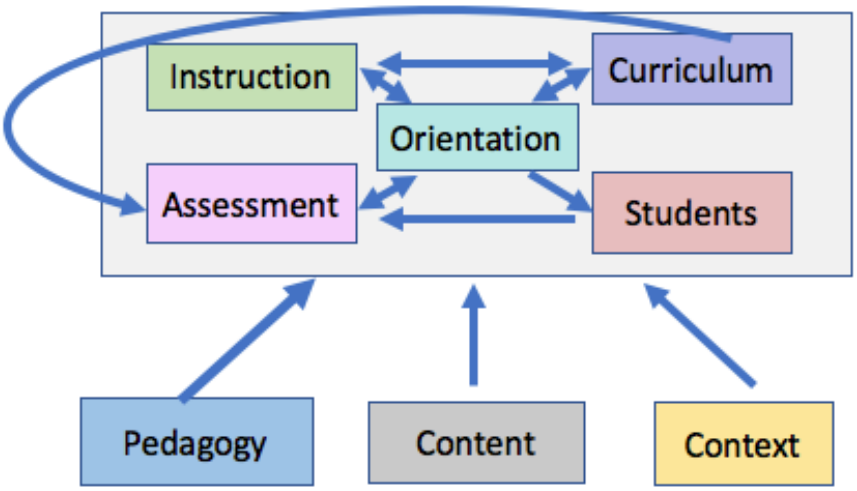
nature of science component of student learning, and considered strategies for doing a follow-up assessment at another point to prompt students to consider that aspect.

In the post interview, Teacher B expressed satisfaction with the assessment, noting that her students as a whole had done well enough that she had not modified instruction for the whole class based on assessment results. She also did not note any changes that she planned to make for future years. She described some of the difficulties that a few students had with the assessment, such as repeating one piece of evidence using different wording rather than stating multiple pieces of different evidence. And she described providing feedback to those students individually in order to support their understanding of the evidence supporting the theory.

6.6. Summary

Teacher B's knowledge of content and of curriculum guided her overall assessment practice in this case. A summary of the knowledge elements that Teacher B referenced and how some of those were referenced, is provided in Figure 5. Arrows in the figure represent elements informing or influencing each other.

Figure 5. Knowledge Referenced by Teacher B



Prior to selecting the assessment, Teacher B had already developed an end-of-unit summative assessment that aligned with state standards and had worked through the instructional materials, adding new material as needed in order to provide students with the knowledge and practice they needed to do well on the summative assessment. Her orientation, of supporting students' learning through a developmental process of building understanding over time, extended to her use of assessment in this case. The assessment Teacher B selected was aligned with her overall goals for the unit, and her goals for the current year were the same as they had been for the prior year, during which she had also done standards-aligned instruction of the same topic, with different materials. Teacher B demonstrated a knowledge of the supports students needed in order to build understanding of the specific topic that she was seeking to assess and intended to use the assessment to inform herself about her students' learning. In part, she intended to use the assessment to see whether the new instructional strategies promoted by the new materials being piloted, were as effective in supporting students as her prior instructional strategies. Her intention was to use the information from the

assessment to guide her in providing additional instruction to support student learning, as warranted by the assessment results.

Teacher B considered the assessment cycle to be successful. From the assessment, she was able to gain an individual understanding of which students understood the evidence supporting the theory of continental drift, and which needed additional instruction. She used this learning to support her students in preparing for their unit test.

In this case, Teacher B's goals and methods of assessment were aligned. The methods used elicited student responses that provided opportunities for insight into the areas of learning that Teacher B was curious about. She used these responses to inform her feedback to students and to inform her instructional decisions moving forward.

CHAPTER 7

DISCUSSION

Several findings become evident through comparison of cases 1 and 2. These are briefly summarized below in Table 6, in terms of features that were similar and different across the two cases and in Table 7, in terms of the topics that are the subject of the research questions. The findings are discussed in detail below.

Table 6. Comparison of Cases 1 and 2, Similarities and Differences

Similarities	Differences
In both cases, Knowledge of Students and Knowledge of Instruction were the intended targets of the assessment.	In Case 1, goals and enactment did not align, and the teacher gained Knowledge of Assessment In Case 2, goals and enactment aligned, and the teacher gained Knowledge of Students and Knowledge of Instruction
The context of piloting a new set of materials as part of an improvement community raised a similar question for both teachers and both teachers sought to answer the question: "will these new instructional strategies bring about adequate learning?" through the selected assessment. In both cases, teachers wanted to compare student learning with new materials to student learning they had previously observed within the same topics, while using other materials.	The teachers responded differently to the context. In Case 1, the teacher's concerns about student difficulties and inadequate instruction with the new materials, played a prominent role in question selection. In Case 2, the teacher's consideration of alignment to standards, preparing students for summative assessment of specific content, and requirements for building student understanding of content, were more salient.
The general strategy for assessment selection in a new context was similar. Teachers selected assessment items that were pedagogically familiar to them.	The enactment of assessment selection was different. In Case 1, a mismatch between the prompt used for assessment and the teacher's goals for the assessment was not caught until after the assessment was enacted. In Case 2, a mismatch between the prompt and the teacher's goals was caught prior to enactment.

Table 6 continued

Orientation played a multi-faceted shaping role in assessment decisions in both cases.	The specific orientations and pathways of influence were different. In Case 1, the teacher's concern about misconceptions and concern that students would not learn key content from the new material was salient. In Case 2, the teacher saw state standards as a guide for setting goals for student learning and, drawing on support from her school culture, prioritized establishing an aligned instructional path for building student knowledge in order to reach those goals.

Table 7. Comparison of Cases 1 and 2, by Topics Related to the Research Questions

	Case 1	Case 2
Goals	Assess student learning and instruction in order to tailor instruction to address a misconception	Assess student learning and instruction in order to tailor instruction to prepare for a unit test
Methods	Analysis question, initially without scaffolding. Decision made during instruction to provide students with additional prompting to elicit responses	Analysis question, initially without scaffolding. Decision made prior to students completing the assessment to provide students with a writing frame to guide responses
Reflections on Student Responses	Student responses were not aligned with the goals of assessment; students provided stories, included some key information in response to additional prompting, but the teacher felt that the assessment was not a true evaluation of student learning	Student responses were aligned with the goals of assessment; students provided the name of a theory, the name of the person who developed the theory, and pieces of evidence to support the theory. Students who had difficulties providing evidence, received additional guidance – but most of the class was able to complete the assessment to the teacher's satisfaction

Table 7 continued

References to Professional Knowledge	Reference to knowledge that significantly influenced the teacher's decision about what to assess included knowledge of instruction and knowledge of students, as well as considerations related to orientation. Reference to knowledge that significantly influenced the teacher's decision about how to assess, was primarily to knowledge of pedagogy and general strategies for assessing student learning	Reference to knowledge that significantly influenced the teacher's decision about what to assess included knowledge of curriculum and knowledge of content, as well as considerations related to orientation. Reference to knowledge that significantly influenced the teacher's decision about how to assess, was primarily to knowledge of pedagogy and general strategies for assessing student learning. The question of aligning methods of assessment to the teacher's goals of assessment was explicitly addressed multiple times during the pre-instruction interview

7.1. Findings Regarding Goals About What to Assess

Our first research question was: How did teachers describe and reflect on their goals about what to assess and the factors they considered to be important in deciding what to assess?

Our findings with regard to this question are:

Finding 1a. Both teachers described goals for assessment that related to eliciting information from students to assess student learning and aspects of instruction for the purpose of informing instruction going forward. This finding aligns with findings from other studies, including Henze, Van Driel, and Verloop (2008) that building Knowledge of Students and, through student responses, Knowledge of Instruction, is a target of assessment. It also aligns with the definition of formative assessment, as assessment that informs teaching and learning – in this case,

through providing teachers with knowledge about student thinking after instruction. A challenge in assessing learning through assessments administered after teaching was shown most clearly in Case 1, where Teacher A found that the students' responses did not provide her with the insight she had looked for into what they had learned from the activity. Her reflections on the reasons for the mismatch between the responses she got and the ones she had hoped for, showed that the missing piece was for students to make a clear statement regarding the fact that she was trying to assess, and then to provide meaningful evidence from the activity to support their answers. The statement of fact would have provided the teacher with knowledge about whether students thinking was correct or incorrect; the evidence would have provided the teacher with knowledge about the connections students were making to the activity they completed in class. Both of these were needed in order to accomplish the teachers' goals for the assessment. Similarly, in Case 2, Teacher B looked for a clear statement regarding what the theory of continental drift was, followed by evidence taken from the prior activities to show that students understood the connections between the activities they had completed and the important ideas that the activities were about. If students had one element without the other, it did not meet the teacher's goals. Use of scaffolding (the writing frame) made it easy for the teacher to interpret students' responses, because the frame specifically prompted students to provide the elements she was looking for. If they did not provide them or provided them incorrectly, the teacher provided feedback to the students. In contrast, the free response format used in Case 1 made it impossible for the teacher to know why students' responses lacked a claim or lacked evidence. In the post-interview, she talked through different options for responses that lacked a clear response, including that students knew the material, but didn't know what was important to address in response to the question; or that they didn't know the material.

This discussion leads to Finding 1b, that aligning methods of assessment with goals of assessment is not automatic, even when goals are specifically connected to a method. An analysis question – like the question in Case 1 – may be created to meet certain goals and for the purpose of eliciting specific information from students; but it may not do so. Teachers' decisions about methods, including how to scaffold assessment and whether the methods selected truly align with the goals that the teacher has in mind, may affect whether goals of assessment are met. So, it is not the case that assessment always informs teachers about students and about instruction, even when that is the goal. How assessment is implemented and the methods selected for implementation are also important.

This leads to Finding 1c, that some kinds of professional knowledge and beliefs may be helpful for teachers in supporting alignment between methods and goals. In Case 2, Teacher B specifically considered alignment between her methods of assessment and her goals for assessment multiple times during the pre-interview. Her general approach to assessment, across all interviews, was to specifically consider alignment of her goals and methods for assessment at multiple levels. She considered the alignment of her unit test (a method of assessment), down to the wording of questions and specific requests made of students, with the standards she wanted students to reach and with the instruction that she intended to provide. She also considered the alignment of her formative assessments (as methods of assessment) to her unit tests and saw the formative assessments as check points or steps along a path. This way of thinking about assessment may have contributed to Teacher B recognizing and correcting a potential mismatch between the formative assessment she selected, and the goals she had for students.

The extent to which the alignment of the goals and methods of assessment in Case 2 can be attributed to the teacher's knowledge about standards, beliefs about learning as a step-by-step process leading to a goal, and knowledge about how to align goals and methods of assessment, is beyond the scope of this study. But, knowledge (1) that alignment of goals and methods is important and (2) of how to align goals and methods of assessment, may be useful to include in the definition of Knowledge of Assessment or in the assessment literacy framework. Additional study would also be needed to understand whether, for teachers who are experienced in considering the alignment of goals and methods, alignment becomes a habitual part of teachers' practices that might be studied through a lens such as situated cognition (Schliemann, 1998).

Below is a summary of findings and implications for the research question: How did teachers describe and reflect on their goals about what to assess and the factors they considered to be important in deciding what to assess?

7.1.1. Summary of Findings and Implications for Question 1

Finding 1a. Both teachers described goals for assessment that related to eliciting information from students to assess student learning and aspects of instruction for the purpose of informing instruction going forward. But, these goals were only met in one of the two cases. Alignment or misalignment between the goals and methods of assessment played an important role in whether or not goals were met.

Implication 1a. Formative assessment can be a key pathway for learning about students and instruction (as suggested by other studies), but is not guaranteed to do so. For formative

assessment to inform instruction in the ways that a teacher intends, goals and methods of assessment need to be aligned.

Finding 1b. Aligning methods of assessment with goals of assessment can be challenging.

Implication 1b. Supporting teachers in practicing and reflecting on how to align goals and methods of assessment may be helpful in increasing teachers' abilities to gather useful information through formative assessment.

Finding 1c. Some kinds of professional knowledge and beliefs may be helpful for teachers in supporting alignment between methods and goals. Specifically, knowledge of standards or curricular goals and how to align assessment to standards as well as knowledge (possibly procedural knowledge) of how to check alignment between assessment goals and methods of assessment may be an area of knowledge worth developing.

Implication 1c. Further study into the professional knowledge and beliefs that support successful assessment practices would be valuable. Also, further investigation of the knowledge supporting alignment between goals and methods of assessment, and teachers' practices for accomplishing this alignment, would be useful. It may be that the definition of Knowledge of Assessment should include knowledge that supports teachers in checking and considering the alignment of goals and methods of assessment. It also may be that the assessment literacy framework would benefit from including alignment of goals and methods as an important aspect of assessment. The types of thinking that teachers do while considering alignment of goals and methods might fit within the "compromises in assessment decision-making and

action-taking" segment of the assessment literacy pyramid proposed by Xu and Brown (2016); but this element of assessment is not currently explicitly addressed.

7.2. Findings Regarding Plans for Methods of Assessment

Our second research question was: How did teachers describe and reflect on their plans about how to assess and the factors they considered to be important in deciding how to assess?

Our findings with regard to this question are:

Finding 2a. Both teachers noted that their familiarity with or use of questions with similar wording in the past, was a reason for selecting the questions they chose.

Finding 2b. Both teachers made adjustments after the pre-interview related to providing increased scaffolding to support students in answering the analysis questions, though the scaffolding selected was different in the two cases

Finding 2c. In Case 1, Teacher A found that her methods for assessing students fell short, in terms of accomplishing her goals and reflected on ways to improve her methods to better align with goals in the future. In Case 2, Teacher B found that her methods for assessing students aligned with her goals, and did not reflect on ways to improve her methods in the future.

Implications with regard to this research question are:

Implication 2a. If teachers select assessment methods based on familiarity, providing both pre-service and in-service teachers with experience in using multiple types of assessment may be an important element in supporting teachers in diversifying their assessment practices.

Implication 2b. Implementing assessment goes beyond a need to select a method. Teachers also make multiple decisions while planning and in the moment, regarding how to scaffold assessments for their students. For these decisions to be helpful in supporting teachers' goals of assessment, reflection and practice may be helpful.

Implication 2c. Instances of unsuccessful alignment between goals and methods may provide motivation and fruitful opportunities for teachers to reflect on and improve their assessment practices. Providing teachers with support through a professional community so that there are opportunities to reflect and discuss, can provide valuable opportunities for learning. Teachers in this study, especially those without a colleague in their school with whom to discuss science assessment decisions, noted that the opportunities for reflection provided through this study were useful to them. Teachers may also benefit from reflecting on instances of successful alignment, though the motivation to do so may be less strong than with instances of unsuccessful alignment.

7.3. Findings Regarding Reflections on Student Feedback

Our third research question was: How did teachers describe and reflect on what they considered to be important about the student feedback that was elicited through assessment?

Our findings with regard to this question are:

Finding 3. Student feedback was considered to be useful when students' responses to the assessment prompts explicitly addressed the internal questions that the teachers were asking (the questions they were asking themselves) about student thinking as it related to instruction. In Case 1, Teacher A described the feedback as being unclear and not aligned with her goals for the assessment, when student responses did not explicitly address the part of the assessment

that was most important to her. She wanted to know whether students believed that humans and dinosaurs could have lived at the same time, and she wanted to see evidence of thinking about the substance of the activity that students had completed, as it related to this question of whether humans and dinosaurs lived at the same time. But, students picked up on other aspects of the assessment prompt that was provided to them and did not address what Teacher A considered to be important. In Case 2, Teacher B described the feedback as being clear and useful. For Teacher B, the internal questions (what she was asking herself) was directly rendered as the assessment prompt that was provided to students. She wanted to know if they could name a theory, name the person who came up with the theory, and state three pieces of evidence to support the theory, using the previous activities. The prompt provided to students asked for exactly those elements.

Implication 3. When planning to elicit student feedback, it may be helpful for teachers to consider the questions they have about their students and about instruction, meaning, the questions that they would like to be able to answer using the student feedback they receive. It may be valuable to specifically check prompts to see whether what students are asked to do, is likely to elicit feedback of a type that will be useful for answering the teacher's internal questions. Further study into the knowledge that helps a teacher to predict the types of answers that students may give in response to prompts, may provide insight into additional knowledge that can support assessment practice. And as with previous implications, supporting teachers in reflecting on alignment between the feedback they are looking for and the prompts they are providing to students, may help in developing this area of practice.

7.4. Findings Regarding Reflections on Professional Knowledge and Beliefs

Our fourth research question was: How did teachers reference their professional knowledge and beliefs while describing and reflecting on the decisions they made about assessment?

Our findings with regard to this question are:

Finding 4a. Teachers referenced their professional knowledge and beliefs in multiple and complex ways throughout interviews focused on assessment plans and practices; and the ways in which knowledge and beliefs were references was different across the two cases.

Implication 4a. The knowledge that informs assessment practice is complex, contextual (in that some qualities of knowledge are significantly different across cases and individuals), and involves multiple elements both within and beyond PCK. More study is needed to understand which elements or qualities of knowledge may be most consistent and significant in supporting teachers' assessment practices and to understand why the elements of knowledge that teachers reference (for example, referencing Knowledge of Students most prominently in one assessment and Knowledge of Curriculum most prominently in another) may differ across cases and individuals.

Finding 4b. Teachers' beliefs and orientations were referenced in a way that was consistent with Orientation playing a shaping role in teachers' thinking about assessment. In Case 1, Teacher A referenced Knowledge of Students and Knowledge of Instruction prominently, along with a concern about students having a misconception and possibly receiving insufficient instruction through the new curriculum. In Case 2, Teacher B referenced Knowledge of Curriculum and

Knowledge of Content prominently, along with a view that student learning involved a process of building toward instructional goals. In both instances, the knowledge referenced by the teachers was consistent with core beliefs that they had about what was important to the practice of teaching and about their roles as teachers. This lends support to the central role of Orientation in the Magnusson, Krajcik, and Borko (1999) model of PCK. It also lends support to Xu and Brown's (2016) placement of "Assessor Identity" at the top of their assessment literacy pyramid.

Implication 4b. How teachers think about what is important for students to learn and even how teachers see themselves as educators and as assessors may play a key role in determining which knowledge teachers refer to while making choices about assessment. It is therefore not sufficient for professional learning opportunities to focus on teachers' knowledge base and it may be important for professional learning opportunities for teachers to support teachers in developing beliefs, orientations, and even elements of identity that support strong assessment practice. More study is needed in this area.

Finding 4c. In both cases, teachers referred to considerations related to multiple elements of their professional knowledge and beliefs as they described and reflected on the factors that shaped their goals for what to assess. In order to make sense of the possible relationships between teachers' knowledge and their thinking about goals for what to assess, it was helpful to consider both the types of knowledge that teachers were referring to (such as, Knowledge of Students), and some of the qualities of that knowledge (such as, knowledge that students may have specific misconceptions, or knowledge that students need multiple opportunities to practice when learning to use evidence to support their thinking).

Implication 4c. While it is useful to use Magnusson, Krajcik, and Borko's (1999) model to categorize elements of teachers' knowledge, it is important to recognize that the details of the knowledge within those categories can be qualitatively different across individuals and across cases. These qualitative differences regarding *what* teachers know or believe about students, or *what* teachers know or believe about instruction as well as *how* teachers select from and apply that knowledge, are central to our goal of understanding how knowledge informs assessment practices.

7.5. Summary

Our purpose in conducting this study was to understand how knowledge can contribute to middle school teachers' decisions about what to formatively assess in science learning, and how knowledge can contribute to middle school teachers' selections of methods for formative assessment of science learning. To that end, we have focused on teachers' goals for specific cycles of formative assessment, methods selected for conducting formative assessment, descriptions of and reflections on the student feedback elicited through assessment, and how teachers' referenced their professional knowledge and beliefs. In summary, we have found that knowledge contributes to teachers' thinking about assessment goals and methods in multiple and complex ways. We have found that formative assessment can inform teachers' practices, but that it does not necessarily do so. In order to obtain information that teachers can use for the purposes they set forth when planning assessment, goals and methods of assessment need to align in ways that elicit valuable feedback from students that teachers can use to answer the questions that motivate their use of formative assessment at a specific point in their instruction. Without alignment, the feedback elicited from students may not be useful for informing

teachers' instructional decisions. The relationships between knowledge and practice of assessment are complex. Orientations and beliefs play an important role, and it may be important to consider teachers' identities as assessors, while working to support teachers in improving their practices. And finally, the knowledge teachers bring to assessment is drawn from across multiple categories of knowledge, and the categories of knowledge that teachers consider most strongly while making assessment decisions may differ across cases. While understanding the decisions teachers make, it is useful to look at the knowledge categories (Knowledge of Students, Knowledge of Assessment) as a way to organize a complex system – but to also look at the qualities of knowledge within those categories. Whether specific categories of knowledge are better for guiding assessment decisions than others, or whether specific qualities of knowledge are more useful than others, is a topic for additional study. Based on this study, we recommend further study of two specific areas of knowledge for assessing student learning. These are: (1) knowledge about how to align goals and methods of assessment so that assessment goals are most likely to be met, for example by drawing on (2) knowledge about how to use assessment methods (prompts or scaffolding) to elicit student feedback of a type that can be used to answer questions the teacher wants answered in order to inform instruction at a specific point in instruction.

CHAPTER 8

CONCLUSIONS

Formative assessment is a critical component of classroom practice, but the process of conducting formative assessment is complex and challenging for teachers. This thesis has investigated how teachers' professional knowledge informed formative assessment practices during assessment cycles conducted in middle school science classrooms. Our goals have been to understand how knowledge can contribute to middle school teachers' decisions about what to formatively assess in science learning and to understand how knowledge can contribute to middle school teachers' selections of methods for formative assessment of science learning.

This study was conducted in the context of a Math Science Partnership in a rural state. At the time of the study, the teachers who participated were in the process of piloting new materials for teaching Earth science at the middle school level. In that context, the following research questions were explored through development of two case studies focused on cycles of classroom formative assessment:

1) How did teachers describe and reflect on their goals about what to assess and the factors they considered to be important in deciding what to assess?
2) How did teachers describe and reflect on their plans about how to assess and the factors they considered to be important in deciding how to assess?
3) How did teachers describe and reflect on what they considered to be important about the student feedback that was elicited through assessment?
4) How did teachers reference their professional knowledge and beliefs while describing and reflecting on the decisions they made about assessment?

Data were gathered through multiple cycles of interviews before and after instruction and assessment, classroom observations, and classroom artifacts (copies of the responses that

students gave to the assessments). The case studies developed for this thesis presented evidence that was primarily from the interview data. Analytic methods included multiple rounds of line by line coding, phenomenology, and a comparative case study approach. Magnusson, Krajcik, and Borko's 1999 theoretical framework for teachers' professional knowledge, was used as a framework for analyzing and organizing the research data.

Findings from this research show that the task of formatively assessing student learning is complex, with teachers referencing multiple areas of their professional knowledge during the process of describing and reflecting on their decisions about what and how to assess in the science classroom. Categorizing knowledge using Magnusson, Krajcik, and Borko's (1999) framework was a useful method for organizing our thinking about what teachers know, but the interplay and blending of these categories was evident in how teachers talked about their assessment decisions. In the cases we studied, teachers frequently referenced multiple categories of knowledge as well as one or more beliefs within a single statement. This integrated nature of knowledge made analysis complex.

Teachers' decisions about what to assess can be influenced by knowledge from multiple categories of their professional knowledge, as well as by beliefs. In the cases we studied, the teachers referenced different categories of knowledge as playing a key role in their thinking about what to assess; they also referenced different qualitative aspects of knowledge. For instance, Teacher A referenced Knowledge of Students (a category) and her thinking about a misconception (one quality of her knowledge within that category) as playing a key role in her thinking about what to assess. Teacher B referenced Knowledge of Curriculum (a category) and her thinking about learning as a step by step learning process building toward a goal (one quality of her knowledge within that category). These very different ways of thinking each contributed to a teacher's decision about what to assess, showing that there are at least multiple pathways

teachers can take in thinking about what to assess. But by what mechanisms are teachers' decisions about what to assess shaped? Orientations may play a mediating role, as suggested by Magnusson, Krajcik, and Borko's (1999) model for PCK, in shaping teachers' thinking about which elements of their knowledge base are important to consider in a specific assessment situation. But other frameworks suggested in the literature may also be useful in understanding the contextual nature of teachers' assessment decisions. One example is a model for assessment literacy proposed by Brown and Xu (2016), which places a teacher's identity as an assessor at the pinnacle of an assessment literacy pyramid that has professional knowledge as its base.

This research also shows that to understand teachers' thinking about assessment in our cases, it was necessary to interact with the teachers' statements at a grain size smaller than Magnusson, Krajcik, and Borko's (1999) categories. Beyond thinking about which categories of knowledge teachers' accessed, the qualities of teachers' knowledge within those categories – specific things that they knew (or believed) about students, about instruction, about assessment, or about curriculum – also influenced their thinking about assessment. For example, in the cases presented in this thesis, both teachers considered factors related to knowledge and beliefs about how students learn; but the content of the knowledge that was accessed was very different. In one case, the knowledge accessed related to thinking about misconceptions; common misconceptions that students might have and what kinds of instruction can help students change misconceptions. In the other case, the knowledge accessed related to thinking about student learning as a step-by-step process of building toward a goal. These different kinds of thinking about students and learning have different implications for how teachers' knowledge informs their choices about what and how to assess in the science classroom.

More study is needed to gain insight into whether it is more productive for teachers to access some categories of knowledge over others while making assessment decisions, or whether some specific qualities of knowledge are more useful for implementing skillful formative assessment in the classroom. But two specific areas of knowledge are suggested as being potentially valuable and important, based on the case studies we presented in Chapters 5 and 6. First, the contrast between our two cases showed that alignment between goals and methods of assessment is important and is not automatic. Knowledge about how to align goals and methods of assessment and how to recognize when goals and methods are misaligned, may be an important component of assessment literacy. Our work showed that a mismatch between goals and methods of assessment can leave a teacher without the information that formative assessment is intended to provide about students and instruction. Therefore, this may be an important area of knowledge to support for teachers and pre-service teachers. The second and related area of knowledge that our study showed may warrant attention, is knowledge about how to tailor methods of assessment (including both prompts and scaffolding provided to students) in order to elicit student responses that will help teachers answer the questions they about their students' thinking and about instruction. In Case 1, student responses were elicited, but the responses could not be used to address the questions the teacher had about learning and instruction. Students responded to the prompt they were given, but not to the portion of the prompt that the teacher considered to be most important. As a result, Teacher A moved forward with instructional decisions based on additional assessment using her observations and interactions with students, rather than on the data she had intended to use to inform her instruction. In contrast, student responses in Case 2 provided useful data that Teacher B used to efficiently answer the questions about student learning and about instruction that were most important to her at that time.

In summary, two types of alignment were important in our cases. One was alignment between goals and methods of assessment and the selection of methods that were appropriate to the goals of assessment. The other (very much related) area was alignment between methods of assessment (such as prompts and scaffolding) and the types of data (in terms of student responses) that were appropriate for informing teachers' questions about learning and instruction. The knowledge needed to support the first type of alignment may involve being able to unpack goals and establish and check relationships between the goals of assessment and the methods that will be used to meet the goals. The knowledge needed to support the second type of alignment may involve considering what kinds of data could be used to answer the questions teachers have about students and instruction, and knowledge about how students are likely to respond to specific types of prompts, and what scaffolding or other measures may be helpful in gathering useful data about student learning. More study is needed to gain insight into the knowledge and practices that may best support teachers in selecting methods for gathering data from students to inform the aspects of learning and instruction that they consider to be most important for informing their instructional decisions. The task of ongoing formative assessment calls for teachers to consider what is important to assess and to tailor data collection to meet their goals as an integral and dynamic part of their instructional practices. It is no wonder that teachers find this to be a challenging task (Cisterna & Gotwals, 2018). Providing teachers with opportunities to collaborate with colleagues and to reflect on their assessment decisions, assessment practices, successes and challenges, and the alignment of the different components of the assessment process, may be one way to support teachers in meeting this challenge.

BIBLIOGRAPHY

Abell, S. K. (2007). Research on Science Teacher Knowledge. In S. K. Abell & N. G. Lederman (Eds.), *Handbook of Research on Science Education* (pp. 1105–1150). New York: Routledge.

Abell, S. K., & Siegel, M. A. (2011). Assessment Literacy : What Science Teachers Need to Know and Be Able to Do. In D. Corrigan, J. Dillon, & R. Gunstone (Eds.), *The Professional Knowledge Base of Science Teaching* (pp. 205–221). Springer. https://doi.org/10.1007/978-90-481-3927-9

Atkin, M. J., Black, P., & Coffey, J. (2001). The relationship between formative and summative assessment: In the classroom and beyond. *Committee on Classroom Assessment and the National Science Education Standards, Center for Education, National Research Council*, *2006*(April 24). Retrieved from http://newton.nap.edu/html/classroom_assessment/ch4.html

Avargil, S., Herscovitz, O., & Dori, Y. J. (2012). Teaching Thinking Skills in Context-Based Learning: Teachers' Challenges and Assessment Knowledge. *Journal of Science Education and Technology*, *21*(2), 207–225. https://doi.org/10.1007/s10956-011-9302-7

Ball, D. L., Bass, H., Blunk, M. L., Cole, Y., Delaney, S., Hill, H., … Zopf, D. (2000). Learning Mathematics for Teaching (LMT) Project.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for Learning: Putting it into practice*. New York: Open University Press.

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, *21*(1), 5–31.

Carlsen, W. (1999). Domains of Teacher Knowledge. In J. Gess-Newsome & N. G. Lederman (Eds.), *Examining Pedagogical Content Knowledge* (volume 6). Springer. https://doi.org/10.1007/0-306-47217-1_5

Charmaz, K. (2006). *Constructing Grounded Theory: A Practical Guide through Qualitative Analysis*. London: SAGE Publications.

Cisterna, D., & Gotwals, A. W. (2018). Enactment of Ongoing Formative Assessment : Challenges and Opportunities for Professional Development and Practice. *Journal of Science Teacher Education*, *29*(3), 200–222. https://doi.org/10.1080/1046560X.2018.1432227

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. https://doi.org/10.1002/tea.20440

Cowie, B., & Bell, B. (1999). A Model of Formative Assessment in Science Education. *Assessment in Education: Principles, Policy & Practice*, *6*(1), 101–116. https://doi.org/10.1080/09695949993026

Erickson, F. (2007). Some thoughts on "proximal" formative assessment of student learning. *Yearbook of the National Society for the Study of Education*, 186–216. https://doi.org/10.1111/j.1744-7984.2007.00102.x

Goodnough, K., & Hung, W. (2009). Enhancing pedagogical content knowledge in elementary science. *Teaching Education*, *20*(3), 229–242. https://doi.org/10.1080/10476210802578921

Groenewald, T. (2004). A Phenomenological Research Design Illustrated. *International Journal of Qualitative Methods*, *3*(1), 42–55. https://doi.org/10.1177/160940690400300104

Grossman, P. L. (1990). *The Making of a Teacher: Teacher Knowledge and Teacher Education*. New York: Teachers College Press.

Hanuscin, D. L., Lee, M. H., & Akerson, V. L. (2011). Elementary teachers' pedagogical content knowledge for teaching the nature of science. *Science Education*, *95*(1), 145–167. https://doi.org/10.1002/sce.20404

Henze, I., van Driel, J. H., & Verloop, N. (2008). Development of Experienced Science Teachers' Pedagogical Content Knowledge of Models of the Solar System and the Universe. *International Journal of Science Education*, *30*(10), 1321–1342. https://doi.org/10.1080/09500690802187017

Hill, H. C., Ball, D. L., & Schilling, S. G. (2008). Content Knowledge : Conceptualizing and measuring teachers ' topic-tpecific knowledge of students. *Journal for Research in Mathematics Education*, *39*(4), 372–400.

Kaya, O. N. (2008). The nature of relationships among the components of pedagogical content knowledge of preservice science teachers: "ozone layer depletion" as an example. *International Journal of Science Education*, *30*(1), 1–28. https://doi.org/10.1080/09500690801911326

Kolomuç, A. (2017). Subject-specific science teachers ' views of alternative assessment. In *Asia-Pacific Forum on Science Learning and Teaching* (Vol. 18, pp. 1–17).

Magnusson, S., Krajcik, J., & Borko, H. (1999). Nature , Sources , and Development of Pedagogical Content Knowledge for Science Teaching. In J. Gess-Newsome & N. G. Lederman (Eds.), *Science and Technology Education Library* (Examining, pp. 95–132). Dordrecht, Netherlands: Springer.

McNeely, M. E. (1997). *Guidebook to Examine School Curricula. TIMSS as a Starting Point to Examine Curricula*. Washington D.C.

National Research Council. (2001). *Knowing what Students Know: The Science and Design of Educational Assessment*. (J. W. Pellegrino, N. Chudowsky, & R. Glaser, Eds.). National Academies Press. Retrieved from http://www.nap.edu/catalog/10019.html

Park, S., & Chen, Y. (2012). Mapping Out the Integration of the Components of Pedagogical Content Knowledge (PCK): Examples From High School Biology Classrooms. *Journal of Research in Science Teaching*, *49*(7), 922–941. https://doi.org/10.1002/tea.21022

Park, S., & Oliver, J. S. (2008). Revisiting the Conceptualisation of Pedagogical Content Knowledge (PCK): PCK as a Conceptual Tool to Understand Teachers as Professionals. *Research in Science Education*, *38*, 261–284. https://doi.org/10.1007/s11165-007-9049-6

Pellegrino, J. W., Wilson, M. R., Koenig, J. a, & Beatty, A. S. (2013). *Developing Assessments for the Next Generation Science Standards*. *Developing assessments for the next generation science standards*. https://doi.org/10.17226/18409

Schliemann, A. (1998). Logic of Meanings and Situated Cognition. *Learning and Instruction*, *8*(6), 549–560.

Schoenfeld, A. H. (1998). Toward a Theory of Teaching in Context. *Issues in Education*, *4*(1).

Shulman, L. (1986). Those Who Understand: Knowledge Growth in Teaching. *Educational Researcher*, *15*(2), 4–14. https://doi.org/http://www.jstor.org/stable/1175860

Shulman, L. (1987). Knowledge and Teaching: Foundations of the New Reform. *Harvard Educational Review*, *57*(1), 1–23. https://doi.org/10.1007/SpringerReference_17273

Xu, Y., & Brown, G. T. L. (2016). Teacher assessment literacy in practice : A reconceptualization. *Teaching and Teacher Education*, *58*, 149–162. https://doi.org/10.1016/j.tate.2016.05.010

Yin, R. K. (2017). *Case Study Research and Applications: Design and Methods*. SAGE Publications.

APPENDIX A: REQUEST TO STUDY PARTICIPANTS

Please select or design a written assessment within Unit D for the purpose of learning about what your students know and can do.

The assessment should be something that will be useful for you in your normal teaching practice, and that you will have your students complete during Unit D of SEPUP, before the end-of-unit test.

Please choose the timing, type, format, length, whether or not to assign grades, and any other specific aspects based on what will be most useful to you.

Shirly and I would like to observe and video your classroom while you teach some of the material that relates to your assessment. Once you have an idea of what type of assessment you will be doing, please let me know about the timing of any classes that you think would be especially relevant and useful for us to see.

Once you have selected or designed your assessment, please contact me so we can set up a time to talk about your plans for using the assessment.

I would like to talk with you again after your students complete the assessment, and would also ask you for copies of the completed assessments, including any of your comments if you have made them.

Thank you.

APPENDIX B: ANNOTATED VERSION OF BASELINE INTERVIEW PROTOCOL

K4A Baseline Interview Questions – 12/2/11

This interview is intended to take less than 1 hour

Open the interview:

Thanks for talking with me today – I really appreciate the opportunity to learn from you. If you don't mind, I'd like to record our discussion today so I have an accurate record of what we talk about without having to write it all down right now. As I mentioned last time we talked, I wanted to talk with you today to learn a little bit about your teaching and your experience with SEPUP so far.

Part A - 20 minutes (could move a couple of these questions to the end as well)

Goal: Learn about experiences that have shaped their teaching up to this point (past training, professional development)

1. *I'd like to start by learning a little bit about how you started teaching….*

> *(*if needed, could add additional prompts or a follow-up… how long, what subjects, where, how did you decide to go in to teaching, what did you do before… the point here would not be to get an exhaustive list of all of these things, but to get an idea about length of time spent teaching, initial motivation for teaching, training for teaching, and past experience. I could always ask further questions in other interviews or at the end of the interview if we have time. At most 1-2 follow-ups related to this question, because of time constraints.)

2. *What do you find to be the most rewarding part of teaching?*

3. *What about the most difficult part?*

4. *What experiences do you think have been most useful in preparing you to teach or in improving your teaching?*

> Could also approach this question as advice–seeking…"what kind of training or practice would you recommend for someone (like me) who is planning to become a science teacher?)

5. *What types of things do you find that you learn from your students?*

> Could probe here – "what types of things?" "can you think of an example?" – this question might solicit ideas about formative assessment?

Part B – 20 minutes

Goal: Learn about teachers' current interaction with the SEPUP materials, and reflections on how SEPUP is different from what they did before

6. *How did you get involved in piloting the SEPUP curriculum?*

7. *What do you think of the curriculum so far?*

8. *Is SEPUP similar to what you've taught in previous years?*

> Could probe here – "how is it similar? Are there any differences? How do you think it compares? Can you give an example?"

9. *Do you notice any differences in what your students are learning with SEPUP from what students might learn through other science curricula?*

> (like in past years or in other classrooms that teach similar material) – this may start to get at how they evaluate student learning… could probe – "can you give an example?" "when do you notice that?" "how can you tell?")

10. *What do you think your students are learning about science?*

11. *Are there important areas of student learning that you think are being covered well in SEPUP? What about any that are being missed or not covered well enough?*

> My intention here is to probe, before I ask about assessment, for what the teacher wants their students to know… (this could touch on the horizontal/vertical axes of the frameworks teachers have for student learning…)

12. *How do you prepare for each class?*

13. *Are there any resources that you are finding to be especially useful to you in teaching SEPUP? Are there any that you wish you had?*

Part C – 20 minutes

Goal: Learn about teachers' current assessment practices, reflections on past practices, and current ideas about assessment

13. *Do you find that you assess your students differently with SEPUP than you did in previous years?*

> Do you have any examples of ways you might have done assessment in the past?

14. *When you plan for teaching, how much time do you spend thinking about assessment? How important is it in your teaching?*

15. *How do you decide when to assess your students?*

Could probe here – what are some of your reasons for assessing students? How do you use assessment in your teaching?

16. *What do you think students should be assessed on?* (or, what do you think

should the purposes of classroom assessment be?)

17. *How many times do you assess your students during each unit in SEPUP? Is the amount of assessment you do similar to what you've done in the past?*

18. *What do you think about assessment in SEPUP?*

Could probe here: Is it strong or weak in the curriculum? Are there enough

assessments? Have you needed to create some of your own? How useful are the end of unit tests? How useful are the rubrics? Have you thought about any changes you'd like to see or resources that would be useful for doing assessment?

19. *Can you describe an assessment you used recently with your students?*

Could probe here – when did you give this? Why did you give it at that time? If teacher-created, how did you create this? Is it something you have used before? Was it graded or scored in some way? After the students completed it, what did you do with it? Were you satisfied with how your students did on the assessment?

20. *How do the Maine Learning Results influence how you assess your students?*

21. *When I begin to teach, how much would you advise me to focus on assessment? What do you think I could learn that would help me to use assessment in my classroom? Are there any resources that you would recommend?*

Close the interview:

Thank you – it's such a wonderful opportunity for me to talk with you. Did you have any other thoughts for me, or any things I might have forgotten to ask about?

At the end of the interview, go on to talk about the first assignment to be designed and plan the next time we will be in contact.

APPENDIX C: PRE-INSTRUCTION INTERVIEW PROTOCOL

What is it that you're planning to do?

What will your students do with the assessment?

What subject matter or content will the assessment address?

What do you think is most important within the subject matter or content of the assessment for your students to know?

What do the assessment (questions) assess? [Were there some different kinds of thinking you wanted to assess? What type of thinking were you hoping to get with this question?]

What did you consider in deciding to do this assessment in this way?

As you were choosing the questions, what kinds of things did you consider? How did you decide what was most important? What tradeoffs were there?

Why did you choose these over the other questions you might have chosen?

What's important about this assessment for your students to do?

How will you evaluate student responses?

What are you hoping to learn from student responses?

What are your goals for the assessment?

Compared to other assessments you do (have done), how good would you say this assessment is? What are its good points? What are its points that are not as good?

APPENDIX D: POST INTERVIEW PROTOCOL

Could you describe what you did with these assessments?

What was the overall quality of your students' work?

Was there anything that stood out for you about their work on this assessment?

How did you approach evaluating the assessment? (Is there a rubric, are there things you were looking for?)

What did/will you do with the student responses?

What did you learn from the assessment?

What did this assessment tell you about student thinking?

What else would you like to know about student thinking? Is there information you would like to get from assessment that you did not get from this assessment?

What did your students learn from this activity? How do you know (what evidence do you see of student learning?)

What did you notice as you evaluated the students work? (Were there things that surprised you? Did students seem to understand the concepts? Did they use evidence? Were you satisfied with their responses – and why or why not?)

How would you rate this assessment (compared to others in the curriculum or others you have done in the past) – how good of an assessment is it?

Will you do anything different because of what you learned by using this assessment? Is there anything from this assessment that you plan to address in class? How?

How do you think completing this assessment benefitted or will benefit your students' learning?

What aspects of student thinking did this assess? What types of student thinking did this assess?

If you were creating the materials (the book), would you ask the question differently – or are there other questions you think the book should be asking?

What were your goals for this assessment? In what ways were your goals met or not fully met?

How do you think the teaching in the lesson related to the goals of your assessment? Did you change your teaching in any way because you were planning to use this assessment?

BIOGRAPHY OF THE AUTHOR

Before joining the RiSE Center as a Master of Science in Teaching student in 2011, Laura Millay grew up in rural Maine, travelled and studied abroad, led trail construction on the Appalachian Trail and hiked 2,000 miles of the Pacific Crest Trail, ran an organic farm business, founded two non-profit organizations and graduated with a B.A. in Development Studies from Brown University. Laura spent two years in rural Thailand, one as an exchange student in high school and again as an exchange student and intern in college. During her second year in Thailand in 2001, she came to appreciate the power and the potential of individuals working together to bring about positive change. She also came to miss Maine's seasons and the benefits of belonging in a rural place. Currently, Laura coordinates education research and project evaluation at the Maine Center for Research in STEM Education (RiSE Center) and lives in Bangor with her two sons. She continues to love hiking, gardening, working outside in all weather and seasons, travel, and adventure – and also deeply appreciates good friends, good neighbors, her amazing co-workers and colleagues, and her home and family. Laura is a candidate for the Master of Science in Teaching degree from the University of Maine in December 2018.