TEACHERS' PEDAGOGICAL DESIGN CAPACITY FOR EMBEDDED FORMATIVE ASSESSMENT:

A CASE STUDY OF MASTER TEACHERS IMPLEMENTING A NEW PRACTICE

by

Andrew Weatherhead

B. A. in International Affairs, University of Maine, 2006

M. A. T. in Elementary Education, University of Maine, 2007

C. A. S. in Instructional Technology, University of Maine, 2011

Submitted to the Graduate Faculty of

the School of Education in partial fulfillment

of the requirements for the degree of

Doctor of Education

University of Pittsburgh

2018

UNIVERSITY OF PITTSBURGH

SCHOOL OF EDUCATION

This dissertation was presented

by

Andrew Weatherhead

It was defended on

July 30th, 2018

and approved by

Kari Kokka, Assistant Professor, Instruction and Learning

Jeffrey Suzik, Associate Professor, Fanny Edel Falk Laboratory School

Dissertation Advisor: Ellen Ansell, Associate Professor, Instruction and Learning

TEACHERS' PEDAGOGICAL DESIGN CAPACITY FOR EMBEDDED FORMATIVE ASSESSMENT:

A CASE STUDY OF MASTER TEACHERS IMPLEMENTING A NEW PRACTICE

Andrew Weatherhead, EdD

University of Pittsburgh, 2018

This study is concerned with how teachers use data about students' ideas in science, elicited from embedded formative assessment probes, to make decisions about next steps in instruction. Since Black and Wiliam's (1998) seminal work on formative assessment, researchers and practitioners alike concerned themselves with formative assessment and its potential benefits to instruction and student achievement. However, the literature shows that a significant hindrance for educators is their ability to plan next steps in instruction in response to data elicited from formative assessment. This study drew upon the concept of pedagogical design capacity and the Design Capacity for Enactment framework (Brown, 2002) as a means to better understand what resources teachers both identify and take advantage of when making decisions about formative assessment outcomes.

The study took a qualitative approach and employed a case study methodology. I engaged two fourth grade teachers at an independent laboratory school in Pittsburgh. These teachers implemented two different formative assessment probes twice. The study collected data from five sources. Observations with video and audio recordings captured teachers working with information elicited from probes in planning sessions as well as planning documents dedicated to science. Teachers annotated unit and lesson plans. Lastly, teachers were interviewed individually multiple times throughout the study and at the end. Analysis of the data collected utilized analytic memo writing and first and second cycle coding.

The study found that teachers drew heavily on personal resources (pedagogical content knowledge, subject matter knowledge, and beliefs and goals) when making instructional decisions about next steps in instruction. Teacher response to data from the formative assessment probes became increasingly responsive to student misconceptions as the unit progressed. Early in the animal unit not many changes were made; later, every lesson was adapted or improvised in response to the data.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

To the faculty of the School of Education, the Department of Instruction and Learning, and my dissertation committee thank you for your guidance, wisdom, and support throughout my journey.

To Ellen, there are no words to express my sincere and deep appreciation of your patience and flexibility. You guided and challenged me to think deeply about my teaching and my research. I am a more thoughtful and intentional educator because of you.

To my parents, your love and support know no bounds. You have always been there for me as my life's journey evolves. I am who I am, because of you. Everything I will ever tackle and become will be because of you.

To James, my journey to Pittsburgh brought us together. You have been at my side through the good times and the rough times. You endured my sleepless nights, my lazy days, and were my rock when I needed you the most. You are my best friend and my soulmate. I cannot wait to see where the future takes us.

1.0 INTRODUCTION

"Elementary teachers are the first line of offense in addressing common misconceptions that follow students from elementary grades into middle school, into high school, and even into adulthood."

Page Keeley (2014)

For many educators, including myself, the *A Private Universe* (P. Sadler, Schneps, & Woll, 1989) video was a pivotal and early example of misconceptions. The evidence provided made such a strong case of just how pervasive misconceptions are, even for the smartest of us. The video featured Harvard graduates who struggled to explain what causes the seasons. Their confusion became evident as they drew upon prior knowledge and mental models of the phenomena. This knowledge did not likely come from recent college course work or even their secondary education.

It is quite probable that these misconceptions had remained with students since their intermediate, or even primary, years when they began to develop space-based models of the Earth-Moon-Sun system. These misconceptions often take the form of students envisioning the Earth being farther from the sun in the winter when it is colder and closer to the sun in the summer when it is warmer. This student thinking is consistent with the "More A – More B" intuitive rule which leads students to misunderstand math and science (Stavy & Tirosh, 2000). Left unchallenged by the formal learning environment this thinking will persist. Even when new information, such as the Earth's tilt, is introduced as being the reason for the seasons, students will integrate this new knowledge into existing models. These new insights often result in students imagining that the tilt causes parts of the Earth to be far enough away from the sun to account for the seasons (Sneider, Bar, & Kavanagh, 2011). While the responsibility for addressing these misconceptions lies with all levels of educators, those of us in elementary education hold considerable influence over when and how this first starts to occur.

1.1 PROBLEM AREA

Formative assessment came into the spotlight after the seminal work *Inside the Black Box: Raising Standards Through Classroom Assessment* (Black & Wiliam, 1998). This review of formative assessment literature concluded that for minimal effort one could realize significant results when using formative assessment. The impact of Black and Wiliam's article, along with an additional article on formative assessment published in 1998, is seen in the nearly 2000 citations they have received.

Formative assessment takes on many forms and can vary considerably along a continuum. Heritage (2007) views this continuum in terms of levels of planning and considers the bookends of the continuum to be “unplanned” and “planned." Ruiz-Primo, Furtak, Ayala, Yin, and Shavelson used levels of formality to describe the continuum, informal to formal (2010). The three broad categories that capture the spectrum across both the level of formality and level of planning are On-the-Fly, Planned-for-Interaction, and Embedded (Furtak & Ruiz-

Primo, 2008; Heritage, 2007; Ruiz-Primo et al., 2010). On-the-fly formative assessment is prevalent in the research literature. It is typically grounded in talk and allows teachers to engage students in repeated feedback loops over a relatively short period.

Figure 1. Formative assessment continuum (Shavelson et al., 2008).

At the other end of the spectrum, embedded formative assessment has found more prominence in the professional community literature. In contrast to the spontaneous nature of on-the-fly formative assessments, embedded formative assessments are constructed ahead of time and purposefully placed in instructional sequences during planning. The ready-to-use nature of embedded formative assessments and its ability to “create goal-directed teachable moments” (Shavelson et al., 2008) has made them a highly useful and flexible set of tools, particularly for elementary teachers.

Page Keeley, a retired senior project officer at the Maine Math and Science Alliance and former President of the National Science Teachers Association, has been at the forefront of the formative assessment focus in the professional community. This work has served as a vital link between research and curriculum, instruction, and assessment for practitioners. Her early work with the National Science Education Leadership Association (NSELA) focused on providing educators with a resource that facilitated easy access to the multitude of resources about various topics in science education. The product of this work was her first book *Science Curriculum Topic Study* (Mundry, Keeley, & Landel, 2009) Her well-known series of books on formative assessment stemmed from the work done with *Science Curriculum Topic Study*. Her formative assessment books now numbers close to 15 volumes, providing educators nearly 400 formative assessment probes for embedding into science instruction. The series' first volume *Uncovering Student Ideas in Science* (USIS) (Keeley, Eberle, & Farrin, 2005) laid the groundwork for professional development work that continues today.

Quality science instruction has long been an area of focus and concern in elementary schools. Some significant issues contribute to this, including teacher comfort with disciplinary content and instructional strategies. The National Survey of Science and Mathematics Education (NSSME) is a nationally representative survey of over 10,000 teachers from 2,000 schools. In 1977 the NSSME was administered for the first time. Since then it has been administered four more times; 1985-86, 1993, 2000, and 2012. The most recent survey examined many issues facing science education including, but not limited to, beliefs about teaching and learning, perceptions of preparedness, and instructional practices (Banilower et al., 2013). The NSSME confirms some commonly held beliefs about elementary science instruction and provides a quantitative look at some of the contributing factors.

According to the report (Banilower et al., 2013), only 39% of elementary teachers feel very well prepared to teach science in contrast to 81% feeling very well prepared to teach language arts. Even more, telling is the breakdown of the various science disciplines where only 29% of teachers feel very well prepared to teach life science, 26% earth science, 17% physical science, and 4% engineering. According to the NSSME, 31% of teachers reported using "non-commercially produced materials most of the time" (Banilower et al., 2013). Of those teachers who did use textbooks, 42% reported skipping parts of the textbook and cited "having different activities that work better than the skipped ones" as their primary reason for doing so (Banilower et al., 2013).

The NSSME study shows that teachers make decisions to change planned instruction. Teachers adapt and improvise when curriculum resources are deemed inadequate. Despite this, teachers have exhibited difficulties making decisions about next steps in instruction (Heritage, 2007) following "instructionally tractable" (Wiliam, 2010) formative assessment outcomes. This difficulty, coupled with difficulties surrounding the enactment of quality formative assessment (Furtak et al., 2008; Hondrich, Hertel, Adl-Amini, & Klieme, 2015), illustrates an area of formative assessment use that warrants further examination. It is this area of difficulty that inspires and informs my inquiry described in this document; to better understand the decisions teachers make in response to formative assessment outcomes.

1.2 PROBLEM OF PRACTICE

This study seeks to understand better the decisions that teachers make following the collection of data from an embedded formative assessment probe. Probes are carefully designed to elicit more than just a correct answer. All student response options are grounded in research about commonly held student misconceptions. As a result, the probe provides teachers with information about student conceptions about the given phenomena. Figure 2. is an example of a probe that teachers could potentially embed into a science unit.

Solar Eclipse

People have always been fascinated by solar eclipses. During a solar eclipse, parts of the Earth experience darkness for a brief time during the day. Throughout time, people have had different ideas about what happens during a solar eclipse. Here are some of their ideas:

A One of the nearby planets passes between the Sun and the Earth.

B The Sun passes between the Earth and the Moon.

C The Earth passes between the Sun and the Moon.

D The clouds block out the Sun.

E The Earth's shadow falls on the Sun.

F The Moon's shadow falls on the Earth.

G The Sun shuts off light for a few minutes.

H The Sun moves behind the Earth for a few minutes then comes back again.

Circle the letter of the idea that you think best explains what happens during a solar eclipse. Explain your thinking about solar eclipses.

Figure 2. Formative assessment probe example (Keeley & Sneider, 2012).

The literature shows that when teachers implement formative assessment, they have difficulties with designing next steps in instruction (Heritage, 2007). Formative assessment must inform instruction and results in changes, or else it is only diagnostic (Black & Wiliam, 2009; Keeley et al., 2005). These variations can take the form of changes to planned instruction or continue with instruction as planned given that the information gathered from the formative assessment indicates that it is the best option. Probes intended to reveal student misconceptions; naïve ideas, preconceptions, incomplete ideas, commonly held ideas, alternative conceptions, misunderstandings, and facets of understanding (Mundry et al., 2009) provide teachers with robust data on which to base decisions about instructional changes. The data affords teacher the opportunity to adapt or improvise the instructional episode in response to information revealed by the probes. Understanding how teacher craft next steps in instruction following embedding formative assessment frames the overarching question for this study.

Brown's (2002) Design Capacity Enactment (DCE) framework in Figure 3. provides a lens for examining the design choices teachers make and the relationship that exists between teacher and tool. Examining teachers' Pedagogical Design Capacity (PDC) for formative assessment will provide an opportunity to investigate why some teachers with varying backgrounds make similar curricular choices or why teachers with similar backgrounds might make different curricular decisions.

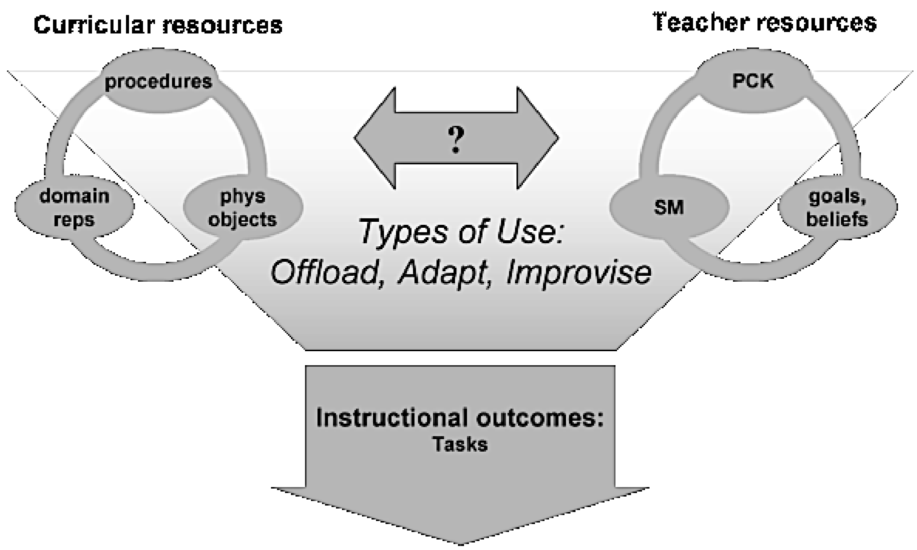


Figure 3. Design Capacity for Enactment (DCE) framework (Brown, 2002)

Brown (2009) suggests that the variations in different teachers' subject matter knowledge (SMK), pedagogical content knowledge (PCK), or certain beliefs may result in offloading; transferring the personal agency of the teacher onto the curricular resources. The DCE framework helps to illustrate how, and potentially why, the same task might be implemented differently by various teachers. Understanding the decisions that teachers make when creating and navigating these diverging paths will help leaders to better support educators engaging in formative assessment.

1.3 INQUIRY QUESTIONS

The following inquiry questions frame this study:

Overarching Question

How do teachers make use of resources available to them to interpret student responses from embedded formative assessment probes and craft next steps in instruction?

Inquiry Question 1

To what extent do teachers offload, adapt, or improvise instruction in response to data collected from embedded formative assessment probes?

Inquiry Question 2

What instructional resources—curriculum, professional development, and other tools—do teachers identify and how do they make use of these resources in response to data collected from embedded formative assessment probes?

Inquiry Question 3

What teacher resources-pedagogical content knowledge, subject matter knowledge, and beliefs-do teachers express and how do they make use of these resources in response to data collected from embedded formative assessment probes?

1.4 DEFINITION OF TERMS

Formative assessment. "Practice in a 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 the decisions they would have taken in the absence of the evidence that was elicited" (Black & Wiliam, 2009).

Formative assessment probe. "A purposefully designed question that reveals more than just an answer. A probe elicits a response that helps teachers identify students' ideas about phenomena or a concept." (Mundry et al., 2009).

Embedded formative assessment. "…comes ready-to-use; teachers or curriculum developers place formal assessments ahead of time in the ongoing curriculum to create goal-directed teachable moments." (Shavelson et al., 2008)

Pedagogical design capacity. "The capacity…to mobilize the curriculum resources in productive ways…perceiving the affordances of the materials and making decisions about how to use them to craft instructional episodes that achieve…goals" (Brown, 2009)

1.5 PERSONAL INTEREST OF THE RESEARCHER

My very first teaching job was replacing a retiring fourth-grade teacher of 35 years. The arrangement of classrooms changed. Consequently, the first impression of my classroom was of 30 years of teaching boxed and piled up in the middle of the room. I dug through the pile, sorting into whatever categories I could discern. When all was said and done, there was a single box of science materials. The instructional materials inside consisted of binders of worksheets, printouts, photocopies, and even some mimeograph pages that had bled through the surrounding pages. It was a curriculum designed over years of teaching. The reasoning for many of the instructional choices would elude me, known only to my predecessor. When I inquired with the other fourth grade teachers, I soon discovered that, apart from a few everyday activities, the unit topics were the only things that were done consistently across the grade level. My induction phase was a time of trial and error as I did my best to adapt and improvise my way through the science units of study; weather, magnets, habitats, and the solar system. During those years, there was a sense of being in survival mode. I needed to get through my units, reflect on what worked and what didn't, adjust my unit plans, and start the process all over again.

My situation is a familiar one to many educators, both new and veteran. While some schools do provide educators with formal curriculum modules or textbooks, many are still using self-designed curriculum and instructional materials. Brown (2009) suggests that "when teachers use curriculum materials to craft instructional episodes to achieve goals when they use materials as tools to transform a classroom episode from an existing state to a desired one, they are engaging in design – whether or not they are aware of it." Indeed, even though I was unaware of Brown's conception of pedagogical design capacity, it is clear that even during those early years of science education I was engaged in the design, limited primarily by my inexperience with my teaching craft.

Enter formative assessment, and what I would argue was one of the most significant transformations of my professional practice in elementary science education. In 2009, I undertook a fellowship with the Maine Governor's Academy for Science and Mathematics Leadership. Page Keeley led the Academy and at the core of the professional development in the academy was her work with formative assessment, *Uncovering Student Ideas in Science,* and *Science Curriculum Topic Study*, her work linking research to curriculum, instruction, and assessment. Through the years of work with Ms. Keeley and formative assessment I have come to appreciate and value the often overlooked, yet vitally important role elementary teachers can play in revealing, challenging, and eventually breaking misconceptions. As Keeley says in her book *What Are They Thinking?: Promoting elementary learning through formative assessment*, “the elementary science classroom is the first line of offense in making sure misrepresentations of science do not shape students’ views” (2014). It is this notion that inspires me to continue working with elementary teachers in formative assessment professional learning opportunities.

2.0 REVIEW OF THE LITERATURE

In this review of the literature, I will discuss the significant focus of this study, formative assessment. I will describe the major themes that resulted from the initial review. I conclude by discussing the pedagogical design capacity the framework used in this study to examine teacher implementation of formative assessment

2.1 FORMATIVE ASSESSMENT

This review utilized ERIC database searches, an analysis of handbooks, Web of Science database searches, and journal article searches. Each review method generated new insights into formative assessment along with additional questions and inquiries. Reviewing the formative assessment literature revealed both common themes and gaps.

2.1.1 Defining formative assessment

While not the most significant finding, attention must first be paid to the definition of formative assessment, or lack thereof. Despite recent critiques of formative assessment (Bennett, 2011) and calls for more significant attention to disciplinary substance (Coffey, Hammer, Levin, & Grant, 2011) the literature lacks a singular definition of formative assessment and more particularly lacks science- specific definitions. Literature reviewed that was specific to formative assessment and science employed a broad, non-disciplinary specific definition (Atkin, Black, & Coffey, 2001; Cowie & Bell, 1999; Falk, 2012; Furtak et al., 2016; Gotwals & Birmingham, 2016; Hondrich et al., 2015; Shavelson et al., 2008).

Table 1. Definitions of Formative Assessment Over Time (1989-2009)

Sadler (1989)	is concerned with how judgements about the quality of student responses can be used to shape and improve the student's competence by short-circuiting the randomness and inefficiency of trial-and-error learning.
Black & Wiliam (1998)	all those activities undertaken by teachers — and by their students in assessing themselves — that provide information to be used as feedback to modify teaching and learning activities. Such assessment becomes formative assessment when the evidence is actually used to adapt the teaching to meet student needs.
Cowie & Bell (1999)	the process used by teachers and students to recognize and respond to student learning in order to enhance that learning, during the learning.
Atkin et al. (2001)	refers to assessments that provide information to students and teachers that is used to improve teaching and learning. These are often informal and ongoing, though they need not be.
Shepard (2006)	a tool for helping to guide student learning as well as to provide information that teachers can use to improve their own instructional practice.
Heritage (2007)	a systematic process to continuously gather evidence about learning. The data are used to identify a student's current level of learning and to adapt lessons to help the student reach the desired learning goal.
Black & Wiliam (2009)	practice in a 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 the decisions they would have taken in the absence of the evidence that was elicited.

The lack of a single accepted definition has promoted numerous variations of definitions from seminal works. Despite the need for its handbook (H. L. Andrade & Cizek, 2010) there is no agreed upon definition. Many works reviewed engaged in the same retelling of the history of formative assessment. Table 1. highlights the evolution of formative assessment since the term first appeared in an attempt to combine and simplify these retellings. The origination of the term by Sadler (1989) focuses mainly on the student actor. The definition from Black & Wiliam's (1998) seminal review of studies expanded the working definition to include both teacher and student and requirement that teaching is adaptive to the evidence collected. Some definitions focused on when formative assessment takes place (Cowie & Bell, 1999), the ongoing nature of formative assessment (Atkin et al., 2001; Heritage, 2007), and the process-oriented nature of formative assessment (Cowie & Bell, 1999; Heritage, 2007). These earlier definitions of formative assessments were much more restrictive and did not allow for my own problem of practice to be well situated within them.

The final definition listed in Table 1 is from Black & Wiliam's recent work to develop a theoretical framework for formative assessment (2009). This definition and the subsequent framework were meant to "define and delimit [formative assessment] within a framework which can also unify the diverse set of practices described as formative" (Black & Wiliam, 2009). Subsequently, they highlight five key points that draw much-needed attention to the broad definition. The first centers on the agent in formative assessment where some studies focus heavily on the student (Cowie & Bell, 1999; Heritage, 2007; Wiliam, Lee, Harrison, & Black, 2004) and others the teacher (Falk, 2012; Furtak et al., 2016; Gotwals & Birmingham, 2016; Hondrich et al., 2015; Shavelson et al., 2008). Black & Wiliam explicate that anyone can be an agent of formative assessment and that each different kind of agent engages in a unique set of processes in the framework. These processes fall into three major categories; where the learner is going, where the learner is right now, and how to get there. Figure 5. illustrates this two-dimensional framework and the subsequent five key points (Black & Wiliam, 2009; Wiliam, 2010).

	Where the learner is going	Where the learner is right now	How to get there
Teacher	**1** Clarifying learning intentions and criteria for success	**2** Engineering effective class-room discussions and other learning tasks that elicit evidence of student understanding	**3** Providing feedback that moves learners forward
Peer	Understanding and sharing learning intentions and criteria for success	**4** Activating students as instructional resources for one another	
Learner	Understanding learning intentions and criteria for success	**5** Activating students as the owners of their own learning	

Figure 4. Aspects of formative assessment (Black & Wiliam, 2009, p. 8).

A single definition of formative assessment, or assessment for learning, does not exist and probably never will exist. This review of the literature has highlighted the wide variability in research. The many facets of formative assessment are best captured in the new Black & Wiliam (2009) framework and definition. The framework proved to be a useful tool as a lens to examine the literature. It also provided a structure that aided in the refining and aligning of inquiry questions to the DCE framework. It is Black and Wiliam's framework that allowed for a broader acceptance of what is a formative assessment. It particularly combats the idea of it needing to occur during instruction. They even illustrate ways that the formative assessment cycle could span days, weeks, months, or even an entire year. Ultimately, it is unlikely that Black & Wiliam's (2009) definition will be accepted as the singular, pinnacle of formative assessment definitions. However, the framework itself does provide an organizational utility that serves as a useful lens through which researchers can view, sort, and filter their respective broad fields of literature.

2.1.2 Research on next steps in instruction

Another critical point highlighted in the Black and Wiliam definition is the focus on the next steps in instruction (2009). This is represented in Figure 4 at the intersection of *teacher* and *how to get there*. One of the stronger studies focusing on the next steps in the instruction was by Cowie & Bell (1999). The study followed ten science teachers in New Zealand for two years and sought to investigate assessment where teachers elicited and used student thinking. Two different types of formative assessment are identified in the study; planned and interactive. These two types are similar to the bookends of the continuum in Figure 1; embedded and on-the-fly respectively. The findings about planned assessment determined that teachers acted on the results of their formative assessments in three ways: science-referenced, student-referenced, and care-referenced. These three ways of next steps of instruction are elaborated upon below.

Despite being over 15 years ago, Cowie and Bell's three categories of response to information gleaned from formative assessment still align well with current efforts in formative assessment (1999). Science-referenced next steps in instruction centers around the idea that once misconceptions have been elicited by the teacher the response to this knowledge is to plan activities directly aligned to the misconception. For example, a teacher who has recently learned that students believe plants get their mass from the soil may choose to challenge this misconception by engaging students in some explorations where plants do not require soil to grow. This strongly resembles the body of work by Keeley providing formative assessment probes designed to reveal research-based misconceptions and provide suggested instructional strategies aligned to the probe (Keeley & Sneider, 2012). These probes and subsequent instructional sequences are most often implemented with the whole class.

The student-referenced response centers on the teacher's awareness of the individual student's changing understanding and providing an appropriate instructional activity in response to this that reflects the following concepts the student is capable of learning. Using the same plant example in the previous paragraph, in this scenario, a teacher may realize that further downstream a student will need to understand that Carbon Dioxide is where the tree gets its mass. Drawing on this progression of ideas, the teacher decides to engage students in explorations that help them understand that air has mass. This may be accomplished by having students weigh inflated and deflated balls. Although the language differs, this strongly resembles the emphasis placed on learning progressions and the awareness they provide. Wiliam describes this interaction the ability to navigate a formative assessment outcome using a learning progression to decide the next step in instruction (2010).

Lastly, was the care-referenced next step which sought to support some aspect of the social interactions between students in response to the information gathered from the assessment. These social interactions mirror our current focus on science practices in the Next Generation Science Standards. For example; a teacher may seek to maintain and enhance student interactions where they were engaging in argument as the next step in instruction.

One primary takeaway from this study was the model that was developed. The model for planned formative assessment was an iterative cycle of eliciting, interpreting and acting upon student thinking. At the center of the cycle was the formative assessment purpose. All three elements of the cycle inform and are informed by purpose. This is a critical take away for my study. As teachers grapple with implementing embedded formative assessment, it will be essential to consider how the purpose of that assessment influences their decision making.

The Cowie and Bell (1999) study align well with the need to engage students in scientific practices outlined in the Next Generation Science Standards (NGSS Lead States, 2013). The practices can be broken into three general categories, investigating, sense-making, and critiquing practices (McNeill, Katsh-Singer, & Pelletier, 2015). The latter, critiquing practices, includes engaging in argument and communicating ideas which by their very nature are a social practice. While students can engage in investigating and sense-making individually, it is often beneficial to engage in these practices with a group and share data from investigations and explanations from sense-making with others (AAAS, 2009).

Other studies have addressed the next steps that teachers take in different ways. Furtak et al. (2016) developed and engaged teachers in a Formative Assessment Design Cycle; explore student ideas, design tasks, practice using tasks, enact the task, and reflect on enactment. The study engaged high school biology teachers over four years and had them reflect upon the enactment of formative assessment at the end of a unit. Their decision making centered around how they would enact the formative assessment differently the following year. After four years of this iterative cycle, the study found that only one element of the design cycle did not show significant increases; task quality. One of the elements that did show significant improvement was the quality of feedback. This cycle is an atypical timeline for the kind of formative assessment we are used to thinking about. However, Wiliam explains that formative assessment does not need to take place during instruction (Cowie & Bell, 1999) but can instead occur outside of instruction over varying periods of time; days, weeks, months, or a year (2010).

Falk (2012) similarly engaged teachers in formative assessment professional development. This took place outside of the regular school year. Teachers engaged in analyzing student work and collaboratively planned changes to instruction. Falk found that teachers tended to draw upon instructional activities with which they were already familiar or existed elsewhere in the curriculum. The struggle to identify next step instructional activities was also a finding in a report from the Center for Research on Evaluation, Standards, and Student Testing. The report was a generalizability study of teachers' pedagogical knowledge. The study found that "using assessment information to plan subsequent instruction tends to be the most difficult task for teachers as compared to other tasks" (Heritage, Kim, Vendlinski, & Herman, 2009). The report notes that the evidence collected from formative assessments lay the groundwork for teachers to take action, but that the evidence itself does not provide the next steps in instruction. This lies solely with the teacher and requires them to draw heavily on the domain and pedagogical knowledge.

This highlights a need for further research in this area. Coffey et al. (2011) argue that too much attention is paid to instructional activities (Furtak et al., 2008; Gotwals & Birmingham, 2016; Hondrich et al., 2015; Shavelson et al., 2008; Wiliam et al., 2004; Y. Yin et al., 2008) and not enough attention is paid to "What is happening in the class, and of that, what does the teacher notice and consider?" In Black & Wiliam's (2009) definition and framework evidence must be "elicited, interpreted, and used...to make decisions about next steps." The heavy emphasis on instructional strategies accounts for eliciting and using the evidence for decision making and accounts for the area that deficits have already been found in (Falk, 2012; Heritage, 2007). These two steps in the definition precede and follow interpreting, which strongly aligns with Coffey et al.'s (2011) call for more research. Attention should be paid to how researchers approach these two areas; separately or as a chronological pair that is inseparable.

When considered through Black & Wiliam's (2009) framework, this lack of consistent research falls squarely within their third strategy, "Providing feedback that moves learners forward." It is a distinctly separate process from the actual design and implementation of formative assessment. Studies that focus on the design, planning, or quality (Furtak et al., 2008; Hondrich et al., 2015) of formative assessment enactment often do not look at subsequent changes in instruction as the primary focus of their research. These studies have shown that the quality of the formative assessment cycle implementation is strongly dependent on teacher awareness about the significance and potential benefits of formative assessment (Ruiz-Primo et al., 2010).

Students and peers engage in multiple phases of the cycle simultaneously through singular processes. Failure to complete the cycle, through lack of feedback, is unique to the teacher. Heritage (2007) highlighted this as an area of struggle, and that formative assessment only works with significant buy-in, and knowledge and understanding of the role formative assessment can play.

2.1.3 Formative assessment is skewed informal along the continuum

One of the reasons that formative assessment is so broadly defined is that it takes on many different variations. Figure 1. illustrates how these variations can be laid out along a continuum characterized by the following categories; the level of planning that goes into the formative assessment, the formality of the formative assessment, and the kind of data sought (Ruiz-Primo et al., 2010).

The three broad categories; On-the-Fly, Planned-for-Interaction, and Embedded, capture the full breadth of formative assessment (Furtak & Ruiz-Primo, 2008; Heritage, 2007; Ruiz-Primo et al., 2010). Examining the literature base in formative assessment revealed a strong skew of the literature to the informal end of the continuum. On-the-fly formative assessment, grounded in talk and discussion, and its ability to allow teachers to engage students in repeated feedback loops over a relatively short period is enticing to researchers, especially given Black & Wiliam's findings that for relatively little effort one could reap great rewards when engaging in the formative assessment feedback loop (1998).

In fact, informal, on-the-fly assessment seems to have become synonymous with the phrase formative assessment itself. This can be seen in the titling of published articles where few dealing with informal formative assessment state so in their titles (Black & Wiliam, 1998; Coffey et al., 2011; Cowie & Bell, 1999). However, the opposite is true for formal, embedded formative assessment. Authors seem to use the "embedded" in titles to distinguish the focus of the published work (Brandon et al., 2008; Furtak et al., 2008; Shavelson et al., 2008; Y. Yin, Tomita, & Shavelson, 2013).

The continuum skew is also highlighted by Coffey et al's critique of formative assessment where all the examples of poor disciplinary substance are informal on-the-fly formative assessments. This type of assessment puts a high demand on teacher's subject matter knowledge (Heritage, 2007; Ruiz-Primo et al., 2010) and we would expect this type of assessment to best highlight teacher knowledge deficits.

Research on formative assessment seems to fall into two major categories; those focused on formal, embedded formative assessment and those focused on informal, on-the-fly assessment that occurs during instruction. The latter appears to be much of the recent research. As stated earlier, this would seem to be due to the ability to engage in iterative cycles of this type of formative assessment within a small period. This also highlights the significant emphasis on formative assessment occurring during instruction. This is important for my research which is firmly rooted in the formal, embedded end of the continuum. An awareness that my research is situated within the lesser end of the continuum emphasizes that many studies may not be generalizable to my situation. This is illustrated well in Coffey et al's (2011) critique of disciplinary substance where all examples draw upon informal formative assessment only giving a few sentences worth of attention to formal, embedded assessment claiming that the same problem exists but is outside the scope of the paper. I would contend that planned formal formative assessment in science provides teachers with stronger tools aligned with learning goals and research on student thinking (Keeley, 2014; Keeley & Sneider, 2012; Mundry et al., 2009). These tools are stronger because they tax teacher subject matter much less than informal formative assessment and allows the teacher to offload their agency into the tool with greater confidence.

2.2 PEDAGOGICAL DESIGN CAPACITY

In his dissertation, *Teaching by Design*, Brown (2002) examined the relationship between teacher resources and curriculum resources, their interactions, and the instructional outcomes of these interactions. Pedagogical design capacity (PDC), a term coined by Brown, represents this relationship. He defines PDC as “teachers' capacity to perceive and mobilize existing resources to craft instructional contexts” (Brown, 2002, p. 70).

Brown identifies three areas that constitute teacher resources: pedagogical content knowledge (PCK), subject matter knowledge (SMK), and beliefs. Other researchers have spent considerable time explaining in greater detail what Brown refers to as teacher resources. Magnusson, Krajcik, and Borko (1999) identified some components of PCK for science teaching. The components are illustrated in Figure 5.

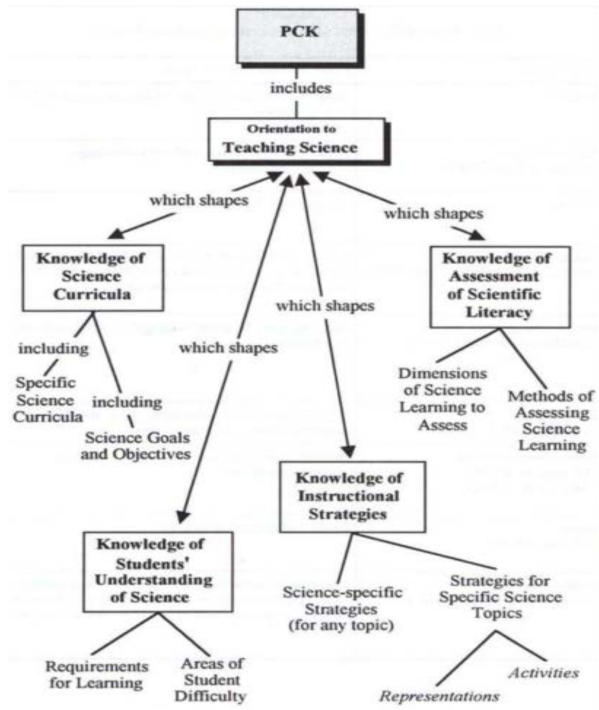


Figure 5. Components of pedagogical content knowledge (Magnusson et al., 1999)

In their discussion of mathematics knowledge for teaching illustrated in Figure 6., Ball, Thames, and Phelps (2008) similarly break down PCK into constituent parts; knowledge of content and students, content and teaching, and content and curriculum. In addition to this, their conception of mathematical knowledge for teaching included another aspect of Brown's (2002) teacher resources, subject matter knowledge.

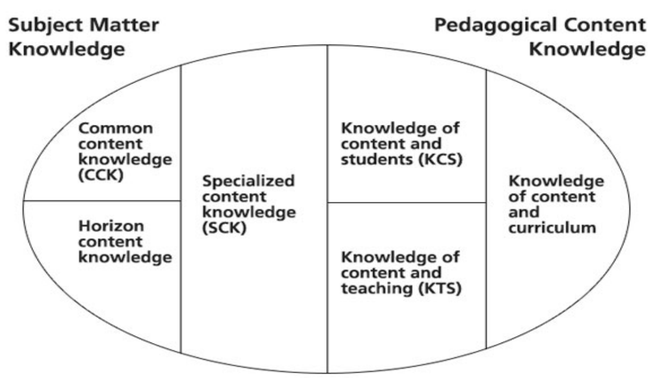


Figure 6. Mathematical knowledge for teaching (Ball et al., 2008)

Luft and Roehrig (2007) developed a set of interview questions about teachers' beliefs in science over the course of five years with pre-service and in-service teachers. Their Teacher Belief Inventory (TBI) elicits teachers beliefs about knowledge and learning in science.

Table 2. contains the five broad categories that capture the belief profiles. Luft also found that these belief profiles aligned strongly with teacher views of science and practices.

Table 2. Teacher Belief Inventory (TBI) categories (Luft & Roehrig, 2007)

Category	Example	View of Science	Practices
Traditional: Focus on information, transmission, structure, or sources.	I am an all-knowing sage. My role is to deliver information.	Science as rule or fact.	
Instructive: Focus on providing experiences, teacher-focus, or teacher decision.	I want to maintain a student focus to minimize disruptions. I want to provide students with experiences in laboratory science (no elaboration).		Traditional
Transitional: Focus on teacher/student relationships, subjective decisions, or affective response.	I want a good rapport with my students, so I do what they like in science. I am responsible to guide students in their development of understanding and process skills.	Science as consistent, connected and objective.	Guided
Responsive: Focus on collaboration, feedback, or knowledge development.	I want to set up my classroom so that students can take charge of their own learning.		
Reform-based: Focus on mediating student knowledge or interactions.	My role is to provide students with experiences in science which allows me to understand their knowledge and how they are making sense of science. My instruction needs to be modified accordingly so that students understand key concepts in science.	Science as a dynamic structure in a social and cultural context.	Inquiry-based

Together these frameworks painted a sharp picture of the knowledge that Brown envisioned that teachers draw upon. However, these personal resources are not the only resources that Brown envisioned teachers drawing upon. Curricular resources, which Brown describes as artifacts, carry equal weight as potential resources. These curricular artifacts are classified as representations of physical objects, tasks, and concepts. The interaction of curriculum and teacher resources, the teacher-tool relationship (Brown, 2002, 2009), is represented along a continuum of potential teacher decisions about next steps in instructions. These decisions can be categorized as offloading, adapting, and improvising.

On one side of the continuum is offloading. Offloading suggests that a teacher "offloads" their agency onto the curriculum materials. Everything is taught just as it was written. The teacher relies on the curriculum materials to support them in teaching. This may be done by a new teacher who needs to rely upon the materials to be able to teach. A veteran teacher may similarly offload their agency onto the curriculum materials but for different reasons.

At the other end of the continuum is improvising where teachers rely solely on their personal resources to create new instructional sequences differ entirely from what was in the curriculum. Brown (2002) cites an example of a teacher who saw value in the discussion between students' competing models of scientific phenomena. The teacher chose to improvise and added a debate over the following week that was not initially in the curriculum.

Adapting is everything that falls in between the improvising and offloading. This may manifest itself in varying degrees, but always keeps part of the originally planned curriculum. The parts that are adapted, or changed, rely more heavily on the teachers' resources.

2.2.1 Design Capacity for Enactment Framework

The three constructs at play described above; curricular resources, teacher resources, and the types of uses come together to form Brown's (2002) Design Capacity for Enactment (DCE) framework illustrated in Figure 7.

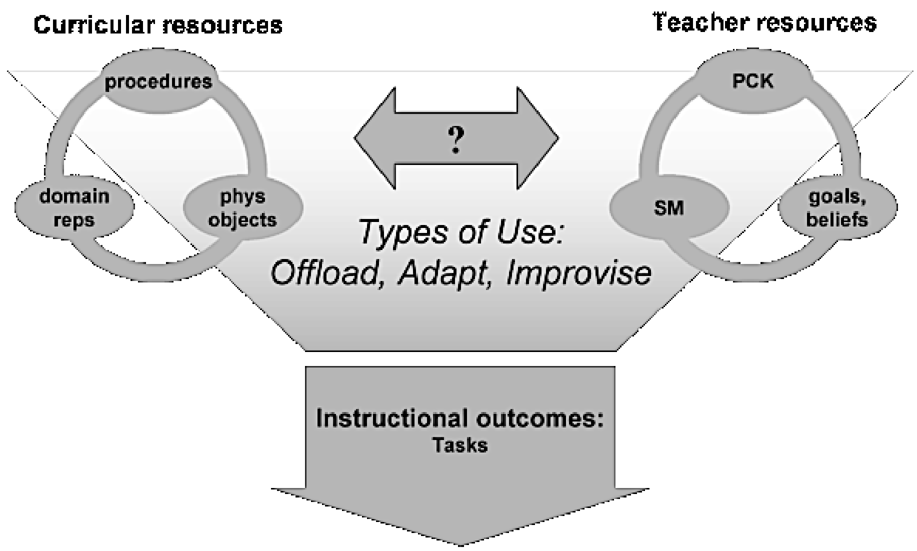


Figure 7. Design Capacity for Enactment framework (Brown, 2002)

The framework has been used in the study of both in-service math (Amador, 2016; Brown, 2009) and science (Brown, 2002; Knight-Bardsley & McNeill, 2016) as well as pre-service teacher education (Arias, Bismack, Davis, & Palincsar, 2016; Beyer & Davis, 2012). Knight-Bardsley and McNeil (2016) recently adapted the DCE framework. In their adaptation, they chose to rename the left-hand side of the framework as instructional resources and added professional development and other tools. Their adaptation of the framework is shown in Figure 8. This broader view of pedagogical design capacity recognizes that teachers may also draw upon recent professional development opportunities when making decisions to offload, adapt, or improvise (Knight-Bardsley & McNeill, 2016).

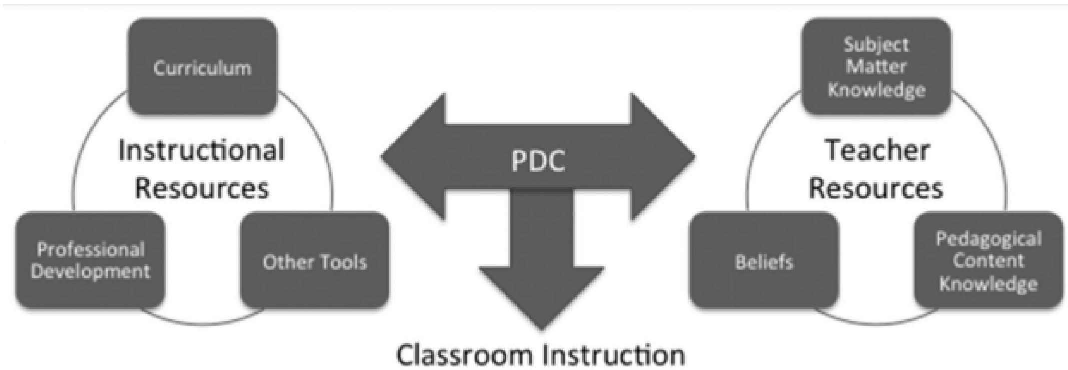


Figure 8. Adapted Design Capacity for Enactment framework (Knight-Bardsley & McNeill, 2016, p. 647)

The adapted DCE framework accommodates the current context of my inquiry study. Despite the master level of a teacher at the site, these teachers are engaging in a new practice and may rely heavily on recent professional development. This adapted framework provides a way to capture this additional dimension of resources that may be drawn upon.

3.0 METHODOLOGY

In this chapter, I set the stage for the inquiry, discussing the setting, stakeholders and the design of the inquiry. I provide a narrative of the methods used and the evidence to collected. I present the types of analyses conducted. This study is framed by the following inquiry questions:

Overarching Question

How do teachers make use of resources available to them to interpret student responses from embedded formative assessment probes and craft next steps in instruction?

Inquiry Question 1

To what extent do teachers offload, adapt, or improvise instruction in response to data collected from embedded formative assessment probes?

Inquiry Question 2

What instructional resources—curriculum, professional development, and other tools—do teachers identify and how do they make use of these resources in response to data collected from embedded formative assessment probes?

Inquiry Question 3

What teacher resources—pedagogical content knowledge, subject matter knowledge, and beliefs—do teachers express and how do they make use of these resources in response to data collected from embedded formative assessment probes?

3.1 INQUIRY SETTING

The setting for my inquiry is an independent school in Pittsburgh. As an independent school, part of their mission is to support scholarly research. It was necessary to be cognizant of the school's unique context as an independent school. For this study, I focused on the area of elementary science. The school has identified elementary science as an area that is in need of improvement. They started to undertake this endeavor through the creation of a part-time science specialist tasked with supporting elementary teachers in a coaching capacity. The school has also started to implement a new plan for professional development in elementary science with a focus on inquiry, formative assessment, and connections to engineering and making. The lack of recent professional development in science created an environment that is ripe for professional learning. The school has not previously spent time focusing on embedding formative assessment in science. The 2017-2018 academic year provided a unique opportunity to explore how best to support teachers in using embedded formative assessment during an academic period coinciding with significant professional development in science.

3.2 STAKEHOLDERS

The critical group of stakeholders is the elementary teachers. These teachers recently began engaging in a year of professional development. They also represent the group from which I selected my case. As master level teachers, they have a significant amount of autonomy in their classrooms. The teachers range in years of experience, science topics taught, and types of curriculums used to teach science. Teacher engagement in formative assessment work mimiced the science specialist's experience; some engaged with formative assessment quickly as early adopters and others took additional time to consider, work with, and adopt the formative assessment strategies.

The science specialist, tasked to work with the elementary teachers, is an individual stakeholder. She has a middle school teaching background and teaches 8th-grade science the other half of her day. She works directly with teachers before, during and after the professional development. Last year I worked with her to support her in her new position. The support usually consisted of identifying potential resources and strategies that are relevant to elementary teachers. She also worked to help teachers during the period when my research takes place. Her passion for the subject has led her to be a voracious consumer of new elementary practices and instructional strategies.

The director of the school is also an individual stakeholder. Two years ago, he piloted the part-time math specialist, a precursor to the science specialist position. He credits the success of the pilot and its ability to overcome hurdles to the approach of the specialist and the fact that the prior adoption of Bridges Math curriculum provided a strong goal for the work. He implemented the science specialist last year because of the success of the math pilot. In addition to the creation of the science position, he perceived a need for future professional development in the area of science. The guiding vision of the professional development is to elevate current science practices and instruction.

The final stakeholder is myself. Outside of my research which will take place this year, I worked with the school to help support the science specialist in her position. I also collaborated with the director and science specialist to develop a professional development plan in science for the current academic year. I was also involved with school as a supervisor for pre-service teachers in two of the university teacher preparation programs. I am also an instructor for students in those programs, teaching their math and science methods courses. My work at the school has allowed me to come to the same conclusions shared by the other individual stakeholders, that elementary science teachers at the school are ready for a year of focused science professional development.

3.3 INQUIRY DESIGN

This inquiry took a qualitative approach which focuses on meaning and understanding the phenomena in question from the perspective of the participants (Merriam & Tisdell, 2016). Of the specific types of qualitative research case study was chosen as the best fit for the inquiry at hand. Stake defines a case study as "the study of particularity and complexity of a single case, coming to understand its activity within important circumstances" (1995, p. xi). The study was a single case study with the animal unit lessons the units of analysis Above all, this was an explanatory case study (R. Yin, 2014) seeking to understand, "How do teachers use data about students' ideas in science, elicited from embedded formative assessment probes, to make decisions about next steps in instruction?"

3.3.1 Overview

The overall study is broken up into three distinct phases; planning, implementation, and end-of-unit. As the names suggest, the planning and end phases are bookends to the major part of this study. The majority of data collection occurred during the implementation phase. Three out of the five data sources in this study are from the implementation phase. This iterative phase will be repeated four times for each teacher, resulting in at least 16 opportunities to collect data. More detailed explanations of the phases follow.

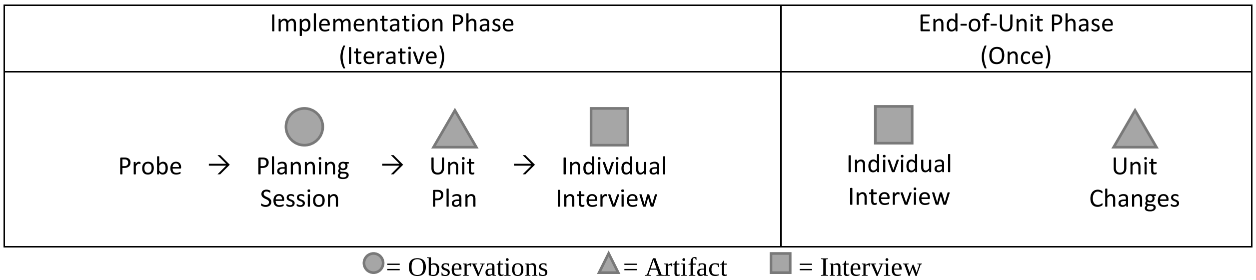


Figure 9. Phases of inquiry design

The first phase depicted in Figure 9. is the implementation phase. This phase spanned the entire length of the science unit and encompassed all occurrences of the embedded formative assessment probes. Teachers were asked to embed two different probes. Each probe was embedded at least twice: an initial assessment and a follow-up assessment. The first probe was embedded a third time for a second follow-up assessment.

Teachers started off by implementing a formative assessment probe and eliciting student responses and reasoning about a given concept. This was captured through the template provided in the formative assessment probe books or by a teacher designated method. Following this the teachers scheduled a time with both myself and the science specialist. This occured during the common planning times built into the new schedule. The primary purpose of these sessions was to allow for teachers to engage with the data about student thinking with the science specialist present as a potential resource. These sessions generated two important sources of data for this study; field observation notes and video recording of the session. These two methods complemented one another well. Video recordings of the sessions allowed for more careful observations to be made after the fact. This was especially useful for times when I was required to take on more of a participant role. Likewise, the field notes allowed for more detailed and nuanced observations of the sessions when I was able to take on the observer role more fully.

Teachers used Google Docs to keep a shared unit plan. Both teachers contributed to this over the course of the unit. This captured changes from the previous year, notes, comments, copies of tasks and other items deemed relevant by the teachers.

At the end of each iteration of the implementation phase a short interview was conducted with each teacher. These interviews were intended to capture aspects of the process and decisions that may be lost if put off until the final phase.

The final phase was the end-of-unit phase. As suggested, this occurred after the unit was completed and when there were no more chances of changes occurring to the unit because of data collected from formative assessment. This phase was characterized by the collection of artifacts and an interview. Artifact collection involved the generation of a teacher created document where teachers characterized the use of each lesson and its connection to the formative assessment probe data. The interviews were individual semi-structured interviews. These interviews were informed by the responses from teachers about any changes made to the unit in response to the data elicited from students.

The data sources outlined above are depicted in the Table 2. matrix and aligned to the inquiry questions. Section 3.4 further explains how each inquiry questions is aligned to multiple data sources and the evidence collected from them. Brief descriptions of the evidence collected by each data source are provided in Table 3.

Table 3. Matrix of data sources and inquiry questions

Data Sources	Inquiry Question 1	Inquiry Question 2	Inquiry Question 3	Overarching Inquiry Question
Observation – Field Notes with Video and Audio Recordings of Planning sessions		X	X	X
Artifact – Unit Plan	X			X
Interview – End of Each Iteration		X	X	X
Interview – End of the Science unit	X	X	X	X

3.3.2 Purposive sample

The factors I elaborate upon in this section make each grade level, classroom, and teacher a unique option for this study. The elementary grades at the school range from kindergarten through fifth grade. I determined that the following characteristics were important considerations when selecting grade level teachers to work with;

- The number of hours per day/days per week science is taught,
- The number of teachers in each grade teaching science,
- The presence of a unit plan with sufficient details; including objectives, procedures, and tasks, and
- More than one teacher having taught at the same grade previously

The fourth grade met all of the criteria and its two teachers were chosen as the sample for the case study. Science is taught in fourth grade five out of six days in the rotation for 45 minutes each day. Both teachers teach science and have taught fourth grade together previously. Their fall science unit lasted the entire fall and was based on a previous teacher created unit.

3.3.3 Participants

Teacher A has a Bachelor of Art in Interdisciplinary Studies and a Master of Arts in Teaching in Elementary Education. She has worked at the school for 19 years since 1998 in various capacities. She began as the Director of After School Programs as well as a resource teacher to both elementary and middle school. In 2001, she became part of the fourth and fifth grade looping classroom. She has taught at these two grade levels for 16 years. She also serves as the team leader for the intermediate team of teachers and serves as a mentor for pre-service teachers.

Teacher B has a Bachelor of Science in Textile Design as well as a Master of Arts in Teaching in Elementary Education. She has worked at the school for 8 years and began in 2009 in a combined age third and fourth grade classroom. She spent two years as a resource teacher and third grade teacher. She moved into the fourth and fifth grade looping position where she has remained. She regularly serves as a mentor to pre-service teachers in various programs.

3.4 INQUIRY METHODS AND EVIDENCE

In this section, I discuss the methods utilized during this inquiry. I detail how observations were used to collect evidence during the planning sessions outlined in the implementation phase. I also explain how evidence was collected using interviews during the end-of-unit phase and the implementation phase. Lastly, I explain how artifacts throughout all phases were used to collect data.

Table 4. illustrates the alignment of methods used with each source of data and the subsequent evidence it elicits. These are in turn broken down by inquiry question. Each method is described in more detail in the next three subsections.

Table 4. Alignment of inquiry questions, methods, and evidence

IQ1 - To what extent do teachers offload, adapt, or improvise instruction in response to data collected from embedded formative assessment probes?	
Methods - Data source	**Evidence to help answer the inquiry question**
Observation - Field Notes with Recordings of Planning sessions	Discussions about changes made about a previous iteration of same probe. Initial decisions about possible ways to act on data collected (Will precede annotated lesson plan and may change before enactment.)
Artifact – Annotated Lesson Plan	Teacher annotated lessons that denote any changes made in response to the probe
Artifact – Unit Plan	Notes about changes made or changes downstream in response to data collected. Notes about changes made about a previous iteration of same probe.
IQ2 - What instructional resources-curriculum, professional development, and other tools-do teachers identify and how do they make use of these resources in response to data collected from embedded formative assessment probes?	
Methods - Data source	**Evidence to help answer the inquiry question**
Observation - Field Notes with Video and Audio Recordings of Planning sessions	Teacher identification and/or use of curriculum resources; existing unit plan, lesson plans, tasks from previous year. Teacher identification and/or use of professional development; engagement of science specialist, use of materials from summer PD
Interview – End of Each Iteration	Responses to questions following up on changes indicated in Lesson Plan Artifacts. Information about returning to resources that teachers were observed using in Planning sessions.
Interview – End of the Science Unit	Responses to questions about patterns noted from all four iterations (i.e. only made use of science specialist once, used teacher supplemental materials from probe every time.
IQ3 - What teacher resources-pedagogical content knowledge, subject matter knowledge, and beliefs-do teachers express and how do they make use of these resources in response to data collected from embedded formative assessment probes?	
Methods - Data source	**Evidence to help answer the inquiry question**
Observation - Field Notes with Recordings of Planning sessions	Teacher surfacing of SMK; common content knowledge relating to animals and/or specialized content knowledge about how teach animals. Teacher surfacing of PCK; knowledge of content and student, teaching, and/or curric. Teacher surfacing of personal goals or beliefs.
Interview – End of Each Iteration Implementation Phase Iteration	Responses to questions following up on impact of probe on teaching practices and why any decisions were made to change instruction.
Interview – End of the Science Unit	Responses to questions about patterns noted from all four iterations (i.e. drew on prior knowledge of misconceptions in all iterations) Response to questions about personal resources (SMK, PCK, Beliefs/Goals)

3.4.1 Observations

As discussed in section 3.1.1 evidence was collected through observations during the planning sessions where the teachers, the science specialist and myself will be present. These planning sessions took place after each implementation of a probe. The teachers engages in examining the student responses and reasoning during these sessions. The science specialist took on the primary role of facilitating this session, providing support to teachers. My own participation took the form of a participant-observer. Yin (2014) notes a number of challenges and opportunities which must be well-thought-out and weighed against one another when considering taking on the participant-observer role. Taking on this role afforded me the ability to manipulate small events. In the context of the planning sessions, being a participant observer allowed me to perceive the experience of the teachers working to make sense of the student responses and reasoning. There is no formulaic correct decision to be made during these sessions. Instead, they are meant to be responsive to student needs, teacher strengths and classroom restraints; ultimately what works best. Being a participant observer allowed me to experience the angst of struggling to understand what we are seeing or the joy from coming to a new-found realization about.

One of the major challenges in this role was the potential for the time spent being an active participant becomes so great that it is detrimental to my ability to be a good observer. While a conscious effort was made to strike an appropriate balance, this was not always possible when the participant demands increase unexpectedly. This threat to credibility was managed through the use of a "checklist of elements likely to be present" from Merriam and Tisdell (2016, p. 140) and the use of video recordings. The video recordings and the subsequent transcripts they generate were used to both compliment my field notes and supplement areas that I missed during the planning sessions. The checklist put together by Marriam and Tisdell (2016, p. 141)can be found in its entirety in Appendix A. Additionally, I made the best effort, immediately after each session, to find a place to complete field notes while the session was still fresh in my mind.

3.4.2 Interviews

During the implementation phase of the study a small semi structured interview was conducted at the end of each iteration of the cycle. This resulted in five small audio recorded interviews with each teacher. These interviews were guided by the interview protocol in Appendix B. The main purpose of these interviews was to capture the impact of the probe on teaching including decisions to make changes and provided an opportunity to probe what else may be happening on days in between observations. The interviews also served as a check on my early analysis of the data as it provided opportunities to clarify my observations of the planning sessions.

The final phase of the study marks the completion of the science unit of study and all possible implementations of the formative assessment probes. During this phase, semi-structured interviews were conducted with each individual teacher. The interviews were conducted as a responsive interview taking advantage of data collected during the implementation phase. I had considerable prior knowledge of the teacher units and collaborative sessions. The core essential questions that framed the interviews are found in Appendix C. The main part of the interview sought to piece together initial patterns observed from previous data collection and ask the interviewee to expand and reflect upon it. This provided both the descriptions and interpretations that Stake (1995) identifies as a primary reason for interviewing. The interviews started off and encouraged a narrative retelling of the unit. Follow-up questions about each session started off more general but led to questions that I saw arise specific to the interviewee. The responsive interviewing also allowed for the interviewee to share and focus on items out of chronological order. This allowed for unique interviewing to occur amongst teachers from the same grade level and is why interviewing was chosen over focus groups.

3.4.3 Artifacts and documents

Artifacts were collected during the implementation and end of unit phases. The primary document collected was the unit plan that teachers generated. This unit plan document was critical in the application of the DCE framework. Decisions to adapt or improvise were represented by actual changes to planned instruction. Multiple changes may take place over the course of a unit and will in turn generate a unique document each time. These materials provide a snapshot of the context and data that frames each observed session. Collecting the physical materials that teachers may draw upon and reference adds to the understanding captured in the video recordings and my field notes.

3.5 ANALYSIS AND INTERPRETATION

In this section, I present how I took on both a deductive and inductive approach. I next describe how I used first cycle codes and coding strategies. Finally, I discuss the strategy I used to analyze data.

3.5.1 Deductive and inductive approach

In this case study, I took both a deductive and inductive approach. The analysis was shaped by the study's conceptual framework, the inquiry questions, and the problem area identified in the literature. These informed the development of a "start list" of codes that were identified prior to the start of data collection in the field and constitute a deductive approach to coding (Miles, Huberman, & Saldana, 2014). The deductive approach was not done at the expense of inductive coding and is in fact complimented by the emergence of codes as the data is collected and analyzed. Merriam and Tisdell (2016) describe qualitative data analysis as becoming progressively more deductive as the study moves from discovering and verifying to testing and confirming when the study becomes primarily deductive. This aligns with the first and second cycle coding approach described below which begins as both deductive and inductive and evolves into a primarily deductive approach.

3.5.2 Analytic memo writing

I engaged in analytic memo writing throughout the course of the study. Saldana elevates the utility of analytic memo writing by highlighting it as a "question-raising, puzzle-piecing, connection-making, strategy-building, problem-solving, answer-generating, rising-above-the-data heuristic" (2016, p. 43). I used the memos to reflect on and about the following opportunities highlighted by Saldana (2016):

- how I personally relate to the participants and/or the phenomenon
- my code choices and their operational definitions
- the participants' routines, rituals, rules, roles, and relationships

- emergent patterns, categories, themes, concepts, and assertions
- the possible networks and processes (links, connections, overlaps, flows) among the codes, patterns, categories, themes, concepts, and assertions
- any problems with the study
- any personal or ethical dilemmas with the study
- tentative answers to my study's research questions

3.5.3 First Cycle Coding

During this first stage of coding the primary goal was to assign the raw data codes through iterative cycles of coding and recoding. The coding methods described below align differently with the deductive and inductive approaches I outlined above. Saldana (2016) identifies over 25 first cycle coding methods in seven categories. The elemental methods category is suggested as a good starting place for the initial coding of data. These methods served as foundations for second cycle coding and were appropriate starting places for a researcher new to qualitative research (Miles et al., 2014; Saldana, 2016).

Saldana (2016) suggests structural coding as a method when working with "start lists" of codes from a deductive coding approach. This approach is often aligned with inquiry questions to allow for an analysis that aids in answering them. This method of coding applies phrases or concepts to the data providing an index that is of use in categorizing the data as well. Saldana (2016) notes that structural coding works best with interview transcripts and open-ended survey questions but not as well with observation field notes.

The Design Capacity Enactment framework in Figure 10. served as the source for structural codes in this study. Each side of the framework provided a top-level code: instructional resources (IR) or teacher resources (TR). Each of these codes were assigned three sub-codes that corresponded to the respective resources. In all, sub-codes represented the six possible types of resources that teachers drew upon. Interviews and planning sessions were transcribed and each turn in the discourse was coded for any of the six resources that were surface or expressed. Codes were also assigned to three instructional outcome options. These codes were applied to the discussion turns as well as the collected unit artifacts. Making decisions about instructional sequences is an inherent piece of the DCE framework. Consequently, discussion turns were also coded for the presence of decisions about instruction. Appendix D contains a full list of codes used.

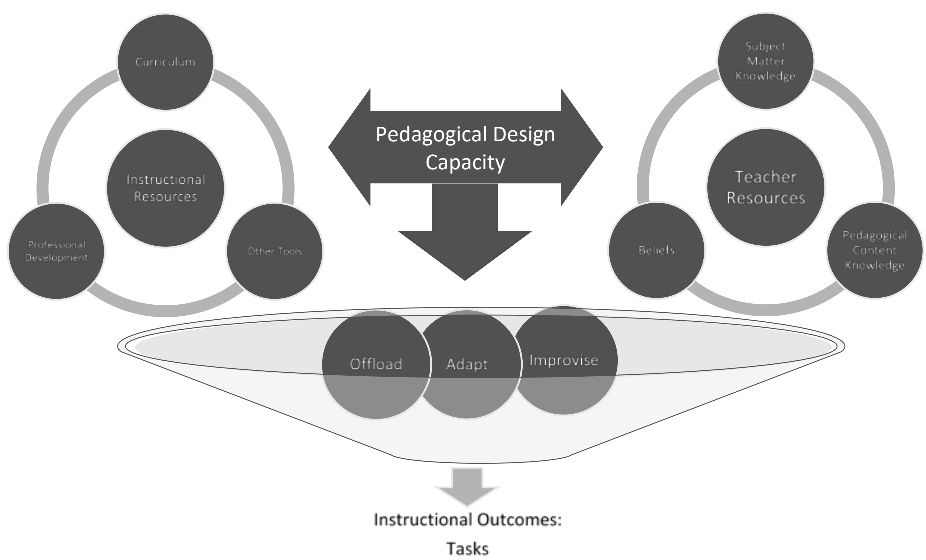


Figure 10. Adapted Design Capacity for Enactment framework (Knight-Bardsley & McNeill, 2016)

Coding beliefs under teacher resources draws on an additional framework to aid in the structural coding of beliefs. In addition to a few questions about probe specific goals and beliefs in the individual interviews at the end of each implementation phase, the interviews at the end of the unit draw upon an additional framework. Luft and Roehrig (2007) constructed a seven question Teacher Belief Interview (TBI) which elicits information about how teachers view learning and knowledge in the science classroom. The responses elicited by these seven question are coded using the five categories: traditional, instructive, transitional, responsive, and reform-based. These questions were piloted with a teacher from the site school. Table 5 illustrates excerpts from the pilot interview that were coded using nVivo.

Table 5. Teacher Belief Interview question pilot

Category	Coded excerpts
Traditional	• sometimes I'm the enforcer
Instructive	• sometimes I'm the instructor • when they convey their understanding, whether it's written or verbally
Transitional	• there's always a range of experience, it's hands on, paper based, observational, discussion. • You can see it for some of them, like their expression • when they're really enjoying • So I think looking for enjoyment, and you can see a lot them when that light bulb goes on • you can tell they're learning because they're actively participating • Some kids you can tell they're learning because they're totally quiet and they're wrapped, you can see that they're just sucking it all in.
Responsive	• "Okay kids, what do you know, what do you want to know?" I mean it's as simple as that. Or, "What do you think you know," that's how I always put it. And sometimes I just will pilot, like, "What are you guys interested in?" • it was a mix of discussion and hands on with a paper pencil component • but then having some kind of written observation journal, or some kind of paper piece.
Reform-based	• Individualize on the fly. Lots of accommodations • it's just a lot of accommodating from work load to flexible grouping, to modifying the written payload of some tasks • And the kid talked for three minutes straight, and I scribed it, so that's kind of how I maximize it. • sometimes I'm just the guide • Then I go back, I look at the standards, make sure I'm hitting those key points, but from there, I can customize and make it responsive to what they're excited about. • some kids aren't on the same mastery schedule as other, so you just have to bear that in mind. And then for the kids that need extra support, that's when you get the parents on deck. But I would say you just look globally at the room. • they're asking questions that are pushing out tangentially from the lesson

3.5.4 Second Cycle Coding

The main purpose of second cycle coding was to work with the data coded in the first cycle in a more inductive approach. This was done through reorganizing the data and looking for large categories or meta-codes that begin to simplify the data. Saldana (2016, p. 233) illustrates the differences between the two cycles with an analogy of furniture that needs to be assembled; first you "inventory all parts" and then "arrange the parts appropriately" before assembling the furniture. While this analogy is concise and to the point, in my opinion it does not capture one of the more important points that there is no one prescriptive way to conduct qualitative data analysis. To this point, the process of working with the data is more analogous to dumping out a new box of Legos. There are many different ways to arrange them before beginning to build and how this is done is highly dependent on the person doing the building. Some people organize by color, some by size or shape, and yet others organize by function or purpose of each piece.

In one way or another, the search for patterns (Miles et al., 2014; Saldana, 2016), pattern matching (R. Yin, 2014), and category and theme construction (Merriam & Tisdell, 2016) are highlighted as one of the more desirable approaches for qualitative data analysis, particularly case study. Pattern coding, as described by Saldana (2016, p. 86), served two primary purposes in this study; to condense the large amounts of coded data generated during the first cycle and to help me in constructing a cognitive map – "an evolving, more integrated schema for understanding local incidents and interactions." The pattern codes came in the form of categories or themes, causes/explanations, relationships among people, and theoretical constructs (Miles et al., 2014).

Recognizing that I am a heavily visual learner it was important to find ways to manipulate the first cycle codes and the emerging pattern codes and visualize them in different ways. To aid in this I utilized computer assisted qualitative data analysis software (CAQDAS). This was used to help in visualizing data in new ways to aid in generating narrative descriptions, matrix displays and network displays. In particular, nVivo provided opportunities look for patterns using compound searches, related code diagrams, and word frequencies and searches.

3.6 CRITERIA FOR QUALITY

In this methodology chapter I have touched upon a number tactics, purposely implemented, to ensure rigor and quality. I used Yin's four design tests to maintain quality (2014). In meeting the first test, construct validity, I have highlighted in section 3.4.1 how multiple sources of evidence will be used. I have laid the foundation for creating chains of evidence during data collection by outlining in Table 3. alignment of inquiry questions with data sources and evidence. I shared early findings report with study participants. In addressing internal validity, I have used triangulation of methods and data sources. I engaged in peer debriefing periodically throughout the study with other members of my STEM cohort. We conducted check-ins with each other every two weeks. To address external validity, I have used purposive sampling in my selection of study participants. I have strongly grounded this study in theory through an extensive review of the literature. Lastly, I have addressed the issue of reliability through triangulation.

4.0 RESULTS

This chapter is organized into four main sections aligned with my three primary inquiry questions and my over-arching inquiry question. Section 4.1 focuses on my first inquiry question setting the stage for the chapter by examining the way teachers used tasks as the unit progressed. Section 4.2 and 4.3 focuses on the resources that teachers identified and made use of. Lastly, in section 4.4 I discuss my overarching inquiry question and how teachers went about crafting the next steps in instruction.

4.1 TEACHER-TOOL INTERACTION

In this section, I address my first inquiry question: "To what extent do teachers offload, adapt, or improvise instruction in response to data collected from embedded formative assessment probes…?" I address the latter half, "…how are these decisions related to the resources they make use of?" after addressing questions one and two. As explained in 2.2, teachers' interactions with the tools available to them can be represented along a continuum and represented by three categories: offloading, adapting, and improvising. At one end of the continuum is offloading where teachers place all of their agency onto the instructional resources. At the other end of the continuum is improvising where teachers put all of their agency on to their personal resources. Adaptation is represented by teachers drawing upon both their personal resources and the instructional resources.

Data about the teacher use of materials—offloading, adapting, and improvising—was gathered at the end of the study through the creation of an artifact. Teachers identified the lessons of the unit that took place during the study and then identified if the lessons were an original part of the unit, adapted from existing lessons, or improvised (brand new). Finally, teachers described whether or not their decisions to adapt or improvise were in response to data collected from one of the formative assessment probes. The lessons depicted in Table 6. that fell within the period of this study spanned three distinct sections of the animal unit. The description is thus divided into the corresponding three sections: no major changes, changing everything, and reviewing and assessing.

Table 6. Overview of animal unit

Months	Unit Topics	Formative Assessment Patterns	Lessons
October	Adaptations	“No major changes.”	**Habitat Change Probe 1**
			Habitats/Survival Needs
			Habitat Rummy
			Bird Skulls Observations/Predictions
			Bird Beak Reflection/discussion
			Bird Beak Simulation
			Piloses Read-Aloud
			Owl Pellets
November			Parent Speaker
		“Changing everything.”	**Habitat Change Probe 2**
	Amphibians		**Is It an Amphibian Probe 1**
			Frog Article
			Frog vs. Toad Venn
			Water-Absorbing Frog Video
			Amphibian vs. Reptile Venn
December			Physical vs. Behavioral Adaptations
		“Reviewing and assessing.”	**Habitat Change Probe 3**
			Is It an Amphibian Probe 2
	Final		FRAMB Vertebrate Graphic

Table 6 continued

	Project		Organizer
January			Create Your Own Amphibian

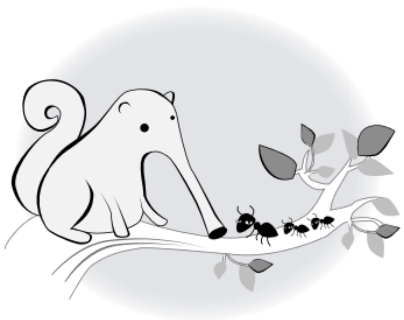
4.1.1 No major changes

The first section of the animal unit focused on adaptations. The activities/tasks for this section mostly focused on birds. Before the start of any instruction for this section, students were administered Probe 1-1, “Habitat Change” (Keeley et al., 2005). The probe, provided as Figure 11, focuses on a fictitious animal that experiences the loss of its food source due to change in the climate. The purpose of the probe was to surface research-based misconceptions that students hold about how animals adapt to deal with changes in their habitat.

Life Science Assessment **Probes** 19

Habitat Change

A small, short-furred, gray animal called a divo lives on an island. This island is the only place on Earth where divos live. The island habitat is warm and provides plenty of the divos' only food—tree ants. The divos live high in the treetops, hidden from predators.



One year the habitat experienced a drastic change that lasted for most of the year. It became very cold and even snowed. All of the ants died. The trees lost their leaves, but plenty of seeds and dried leaves were on the ground.

Circle any of the things you think happened to most of the divos living on the island after their habitat changed.

A The divos' fur grew longer and thicker.

B The divos switched to eating seeds.

C The divos dug holes to live under the leaves or beneath rocks.

D The divos hibernated through the cold period until the habitat was warm again.

E The divos died.

Explain your thinking. How did you decide what effect the change in habitat would have on most of the divos?

Figure 11. Habitat Change probe (Keeley et al., 2005)

Both fourth-grade teachers, the K-5 science specialist, and I were present at this first session. The teachers examined the probes for both the student responses and their reasoning. Instruction for this section continued until early November. I met with both teachers for Individual Interview 1 on the morning of November 9th. These interviews focused on what had been taught since the Planning session 1, a reflection of what had been discussed in the planning session, and beliefs about how students might do on the upcoming probe. Data collected about this span of time on the lessons are listed chronologically in Table 7.

Table 7. Lessons from the first section of the animal unit

Lesson	Related to data from probe?	Offloaded	Adapted	Improvised
Habitat Change Probe 1		X		
Habitats/Survival Needs	No			X
Habitat Rummy	No			X
Bird Skulls Observations/Predictions	No			X
Bird Beak Reflection/discussion	No		X	
Bird Beak Simulation	Yes	X		
Piloses Read-Aloud	Yes			X
Owl Pellets	No	X		
Parent Speaker	No			X

The eight lessons that took place during the first section of the unit fell into all three categories of use within the Design Capacity Enactment (DCE) framework. The first two lessons, *Habitats/Survival Needs,* and *Habitat Rummy,* were brand new to the unit and aligned with the improvise category of the DCE framework. It is important to note that, despite being improvised, the teachers indicated that the decision to incorporate these two lessons into the unit was not informed or related to the data gathered from student responses to the probe. Both teachers shared that these two activities were found in a resource that they already had available to them *Project Wild K-12 Curriculum and Activity Guide*. During an informal discussion when planning a date for the first planning session, Teacher B showed this resource to me and indicated her desire to use the two lessons. The lessons chosen were selected from a part of the *Project Wild* guide that focused on wildlife populations. At that time, she indicated that she had not yet shared the resource with Teacher A, suggesting that these initial lessons were not discussed before the start of the section and consequently were added to the unit plan just before they were taught. Despite the student misconceptions revealed during Planning session 1, the teachers did not make use of the chapter of the guide that focused on changes and adaptations.

The teachers recorded in the shared document that the primary reason for including *Habitat Rummy* was because they “wanted to have the students play a game.” This highlights an early example of how the teachers’ inclusion of lessons is sometimes based on their interest in the engagement of the students. This is consistent with views that were surfaced during final interviews where both teachers emphasized that students learn best when actively engaged in an activity that is hands on. Teacher A also went on to describe her role as a teacher as needing to make sure students are “enjoying it and liking. It is a big part of the hook, getting them interested.”

The Bird Skulls Observations/Predictions were improvised as well and not related to data collected from the *Habitat Change* probe. However, Teacher B shared that the primary reason for adding this to the unit was to incorporate museum loan materials from the Carnegie Museum’s Educational Loan Materials. Similarly, the Bird Beak/Reflection/Discussion was unrelated to the data from the probe however it was adapted from an existing discussion activity about bird beaks that historically preceded the *Bird Beak Simulation*.

Only at the fifth lesson of the adaptations section do we see the first account of the lesson taught being related to the data from the student probe. In this case, the *Bird Beak Simulation,* is a lesson that was already a part of the animal unit. It can be categorized within the DCE framework as a lesson where the teachers offloaded. This is one of only two lessons categorized as offloaded that were carried over from the original unit and not adapted. It is also the only lesson that was offloaded and related to the data collected from the formative assessment probe. This is surfaced in the discussion during Planning session 1 when the teachers grappled with the belief that the Divo could just switch to a new food. The teachers equated this with misconceptions about the adaptability of the mouth and the connection to the concepts addressed in the Bird Beak Simulation.

> **Teacher B:** [Verbalizing the misconception she is typing in the notes] Teeth and the anatomy of the mouth can adapt quickly.
> But really these are adaptations that happen. And that's what the bird thing will address, right? The adaptation of the beaks?
> **Teacher A:** Very slowly. The difference in beaks?
> **Teacher B:** So on this one-- yeah. Bird/beak activity can address this misconception. Digestion can adapt too, right? Can adapt quickly. That's a misconception. It takes time for-- did I already write that?

Teacher A goes on to reflect in her final interview that "because it involved adaptations, we were like, 'This is good. This is another reason why we should do it because it talks about adaptations' we were like, 'Oh, yeah. This is perfect to do and then it will reinforce and help us to teach those things about adaptations.'" While no change was made to the lesson, it is consistent with the definition of formative assessment laid out in Section 1.4 where Black and Wiliam establish that instruction need not change if the decision for the current instruction is "better founded than the decisions they would have taken in the absence of the evidence that was elicited (2009)."

The *Piloses Read Aloud* was an improvised lesson that was directly related to the probe. This came about in a unique way. Instead of the improvised lesson being one that the teachers found or created, the book was sent in by a class parent. The parent had heard from their child that they were studying about adaptations and recommended the book. Teacher A shared that they chose to use the book as a read aloud "because of the probe. We wanted to discuss more about adaptations." The book provided a means to do so in a way that was very similar to the fictitious Divo in the *Habitat Change* probe.

The other lesson categorized as offloaded was the *Owl Pellet* lesson. This lesson was not related to the data from the probe. When asked directly in the final interviews, Teacher A reiterated the lack of any connection to the probe. She identified a few conceptual connections, but the decision to teach the owl pellet lesson was in no way related to the probe. She goes on to share, "we felt like one of the reasons why we like doing that one so much with the bird part is the kids have always, we have found, love getting into the owl pellet, identifying what the thing had eaten, the bird had eaten." This is consistent with reasoning discussed above about the *Project Wild* lessons and the role of the teacher to provide engaging activities that hook the students.

The final lesson of the adaptations section was the *parent speaker*. This was a new improvised lesson that was not related to the data collected from the probe. The parent was a mammologist that the teachers had learned about. Both teachers indicated that they always invite parents in when they have knowledge related to the units they teach. The regularity of this practice speaks to the consistent nature of the parent community at the school.

In summary, this section of the animal unit only had two out of eight lessons that were related to the data collected from the formative assessment probe. Of those two lessons, neither of them were opportunities where the teachers sought out an instructional outcome or task purposefully to address misconceptions revealed by the probes. The lessons either already existed or were presented to them unsolicited. The section also had a few missed opportunities to incorporate more instructional sequences about animal adaptations. The teachers did not adapt existing lessons in response to the data in order to better address the student misconceptions. One could argue that instruction changed very little in response to the data from the probe. Even when a lesson was related to the probe data the lesson was believed to already be adequate to "reinforce and help us to teach those things about adaptations." Overall, this first section is exemplified by the teachers' belief that the instruction that was taking place would adequately address the surfaced misconceptions.

4.1.2 Changing everything

The second administration of the *Habitat Change* probe was completed and originally intended to be the final administration of this probe since the unit would be shifting to a new section. The teachers found that the student data examined in the session revealed that student thinking had not shifted and that many misconceptions still remained around the phenomena of adaptations. In particular, the teachers noticed that the misconceptions fell into two distinct categories: physical and behavioral. Shortly after the planning session, a decision was made to continue addressing adaptations in the next section of the unit because of persistent misconceptions. This entailed the current probe being administered a third time and two different probes used during the Amphibians instruction.

The next topic of the unit focused primarily on Amphibians. According to Teacher A, the focus on a singular class of vertebrate set students up for "our final project, or assessment for this is gonna be the create an amphibian, where the kids will use what they've learned about what is an amphibian to create one." The teachers modified the first probe for this topic, *Is it an Amphibian?,* to only included amphibians and reptiles. It asked students to determine whether each of the eight animals shown was an amphibian and then provide their reasoning for that decision.

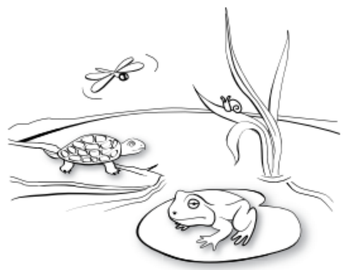
Life and Its Diversity

4

Is It an Amphibian?

Belinda's sister came home excited about her science class. They were studying amphibians. Her sister asked Belinda to help her make a list of amphibians.

Put an X next to the animals they should put on their list.



_____ tree frog	_____ dragonfly	_____ pond turtle
_____ water snake	_____ penguin	_____ beaver
_____ shark	_____ alligator	_____ bullfrog
_____ duck	_____ salamander	_____ whale
_____ catfish	_____ toad	_____ sea turtle
_____ mosquito	_____ crab	_____ seal
_____ snail	_____ rattlesnake	_____ eel

Explain your thinking. What rule or reasoning did you use to decide if something is an amphibian?

Figure 12. Is It an Amphibian? probe (Keeley, 2011)

Teacher B described their reasoning for making these changes in Planning session 2, "So we thought it would be better to also have them isolate specific things and say, 'Well, why is this an amphibian or why isn't it?' That will flush out more of the thinking about what an amphibian is and limit the choices." She went on to share in an Individual Interview that "the probe that was printed in the book, it had lots of different animals. So, I think that the reason that we switched it is, I wanted to hear ... the open-endedness of the probe, it was like, 'Here's 12 different animals. Which ones are amphibians, and why?' I guess I wanted to know more, specifically how they could understand what an amphibian was, and this forced them to explain a few choices, rather than many." In contrast, the original probe shown in Figure 12. (Keeley, 2011) included examples from several different animal classes asking students to put an X next to each animal that was an amphibian and provide reasoning for how they chose the animals they did.

Planning session 3 revealed a number of unexpected misconceptions and surprises surrounding student's difficulties discerning amphibians and reptiles from one another. Data collected from this span of time on the lessons are listed chronologically in Table 8. These were largely informed by what was learned from both probes and planning sessions 2 and 3.

Table 8: Lessons from the second section of the animal unit

Lesson	Related to data from probe?	Offloaded	Adapted	Improvised
Habitat Change Probe 2		X		
Is It an Amphibian Probe 1			X	
Frog Article	Yes			X
Frog vs Toad Venn	Yes			X
Water-Absorbing Frog Video	Yes		X	
Amphibian vs Reptile Venn	Yes			X
Physical vs Behavioral Adaptations	Yes			X

The second section of the animal unit is characterized by the data, from both the *Habitat Change* probe and the *Is It an Amphibian?* probe, informing instructional decisions for all five lessons. This represents a concerted effort by the teachers to address the misconceptions. None of the lessons in this section were original to the unit and as such there were no cases of offloading. The original animal unit focused on Amphibians for a single week. Only two lessons from that week-long amphibian focus were adapted for use this year: *Water-Absorbing Frog Video* and

Create Your Own Amphibian (discussed in Section 4.1.3). The remaining four lessons were all brand new to the unit and fell into the improvise category of the DCE framework. The optimism of the first section gives way in this section to continuing data from the *Habitat Change* probe that student misconceptions continue. In addition to this, the misconceptions that surfaced in the amphibian probe were equally unexpected. Teacher B shared her surprise in the misconceptions revealed by the amphibian probe and eluded to planned instructional sequences that were abandoned due to the basic nature of the misconceptions present. "I think that if everyone had pretty much been able to distinguish the things that we expected ... if kids knew that, yes, salamanders and tadpoles are amphibians, snakes are reptiles, then we probably would have spent more time talking about the more unusual ones, like the Caecilian, and instead, we spent more time on the basics."

The first two lessons of the section, the *Frog Article* and the *Frog vs Toad Venn,* were both brand new, improvised lessons. Both teachers shared that they decided to include these two lessons because they "learned from the probe that students had misconceptions about amphibians."

Similar to the way improvised lessons came about in the first section, the *Frog Article* came as a reading in the ReadingWorks weekly reader and was a chance find. Teacher A shared that, "the kids read this with the student teacher and then she also talked about, we modeled how you can markup important things by underlining or circling on some resources. So, we read this with them, and so they got a little bit more information about what are amphibians." Up until now improvised lessons directly related to the data were from resources that the teachers already had at their disposal or did not have to go and seek out. The *Frog vs Toad Venn* came about in a slightly different manner. This is the first example of where the teachers purposefully sought out a lesson that could help them respond to the misconceptions surfaced by the probe. Teacher A shared that “we found this and it just covered all the criteria, and it just worked perfectly `cause I was like, ‘well we have kids that still don't understand that.’”

The *Water-Absorbing Frog Video* was the first example of an adapted lesson that was informed by the probe data. It is also the first lesson that addresses misconceptions about animal adaptations in the context of amphibians. The previous version of the animal unit contained an amphibian lesson that engaged students with a number of introductory level videos about amphibians from sources such as BrainPop and NatGeo. The structure of the lesson was maintained; however, the actual videos were replaced with the water-absorbing frog. The idea to use this video first surfaced during Planning session 2 when the teachers were discussing the second round of student responses from the *Habitat Change* probe. Teacher B noted that “the amphibians, where they're [students] moving next, like there's a number of ways that we could--like frogs hibernate, right? Don't most amphibians hibernate in the winter? Don't they dig under the mud? There's a cool video that I showed [goes on to talk about the video content].” This is also the first mention of discussing adaptations in the amphibian section of the unit. Teacher B later reflected that they chose to use the video because it “emphasized behavioral AND physical adaptations which we knew from probe needed to be addressed.”

The last two lessons of the section were both related to the findings from the probes and were examples of improvising within the DCE framework. The improvisation that took place for these two lessons is unique within this section, as both of these lessons were created by the teachers. The first lesson compared amphibians and reptiles using a Venn diagram. This format built on a prior task and took advantage of student familiarity with the process. Teacher A shared, “we did another Venn diagram and had them ... gave it more characteristics about what are specific to an amphibian and then what's specific to a reptile because that was another misconception and this definitely we picked out of that, like we were like this is something that would help clear up some of the misconceptions here.” The final lesson of the section addressed the continuing topic of animal adaptations. The students created a T-Chart where they sorted amphibian behavioral and physical adaptations that they learned about through various methods. Teacher A identified this as being a direct need from the probe and that to further help students understand the difference they “explored books and animal fact cards to find other example of adaptations.”

4.1.3 Reviewing and assessing

The third implementation of the *Habitat* Change probe was administered at the end of the amphibian instruction. This third *Habitat Change* probe was modified to ask students to explain their reasoning for each choice as opposed to the original singular reasoning question. This mirrored the modifications that were made to the amphibian probe. Teachers expressed concern that a student could simply select that the Divo died and get away with not having to explain why they though that the other options weren’t likely. The discussions in this session focused largely on individual student conceptions about each response and how they changed over time.

Teachers discussed what was done to address the issues surrounding student conceptions of amphibians and teacher beliefs going into the second and final administration of *Is it an Amphibian.* During this planning session the same group discussed and discovered numerous similar misconceptions about amphibians.

While the second section was characterized by zeroing in on specific misconceptions with focused lessons, this section moves in the opposite direction and flares to encompass broader concepts. This is done in a review capacity prior to the end of the unit project.

Table 9: Lessons from the third section of the animal unit

Lesson	Related to data from probe?	Offloaded	Adapted	Improvised
Habitat Change Probe 3			X	
Is It an Amphibian Probe 2			X	
FRAMB Vertebrate Graphic Organizer	Yes			X
Create Your Own Amphibian	Yes		X	

The two lessons that took place in this section continue with the pattern seen in the last section where all lessons were related to the data from the probes. These last two lessons are unique in that they are informed by the prior administered probes, yet do not have a follow up probe. In these final lessons of the unit there are no offloaded lessons that were already part of the unit.

During Planning session 4, following the *Habitat Change* probe, an impromptu table graphic organizer task was developed in response to what was learned about student thinking. This is in keeping with the last two improvised lessons from section two where the teachers created the lessons themselves. The creation of the FRAMB (Fish, Reptile, Amphibian, Mammal, Bird) Graphic Organizer task during the recorded planning session provided unique insights into how the teachers and the specialist co-develop lessons. During the session, Teacher A noted, "from [the] probe, learned that there were still misconceptions about all animal groups characteristics.” In response to this Teacher A shared that she had originally planned to go back over the material as a whole class and have students take notes on animal characteristics as a review. Teacher B echoed similar reasoning and expressed her own thoughts on what students were taking away from the unit, “I feel like that's probably how I will follow up also because to me it seems like the most important take away that I want them to have and they are not taking that away yet.”

The last lesson of the unit was an adaptation of a lesson from the original unit. As mentioned in Section 4.1.2 this is only the second example of an adapted lesson where the changes were informed by the data from the probe. The lesson was originally designed to immediately follow the frog video lesson. The lesson now serves as a final project of the unit, capturing student conceptions about amphibians, habitats, and animal adaptations. Adaptations were the primary content modifications made to the lesson. The teachers also made the expectations and structure of the lesson clearer incorporating “graphic organizers for universal design best practice” as indicated in the unit lessons artifact generated by the teachers. Teacher B highlights just how the *Create Your Own Amphibian* lesson brings together the two topics that were the focus of the probes,

> “And I think the amphibian activity is bringing it all together--with the adaptations because they have to come up with the adaptation. And over the last couple of days I've been needling them to justify why that adaptation-- I was like, ‘You can't just write that it has sticky fingers.’ Like, ‘What does that have to do—' an adaptation serves a purpose in keeping you alive in your habitat. So what purpose does those serve? They're like, ‘Oh, well then, if they didn't have the sticky fingers they wouldn't be-- they needed those to climb up the bamboo shoots or something like that.’ So I'll say, ‘Good, then write that down.’ So I think here that's been fresh on their mind, right?”

In summary, the teachers displayed noticeably different responses to the data collected following the administration of each probe. Both chronology and the concept addressed provide a means to consider how these responses can be categorized. The response to the very first probe about habitat change was one of “no major changes”. These lessons addressed the content of adaptations and signified a period of confidence in the ability of existing lessons to address the misconceptions surfaced in the probe. This period of “no major changes” also contains the only instance of offloading during the study. This is consistent with the notion that teachers perceived certain affordance of the curriculum when deciding to retain the lesson as is. The introduction of the amphibian probe and the continuation of the habitat change probe signifies a period of significant change where lessons were either adapted or improvised. This radical period of change was linked back to the data from the formative assessment probes and suggests that teachers were approaching the latter half of the unit with a different mindset than the beginning of the study.

4.2 INSTRUCTIONAL RESOURCES

In this section I address my second inquiry question: “What instructional resources—curriculum, professional development, and other tools—do teachers identify and how do they make use of these resources in response to data collected from embedded formative assessment probes?” Instructional resources are broadly defined as the resources external to personal resources and are divided into the three categories: curriculum, professional development, and other tools. Analysis of coding revealed that the case study teachers identified instructional resources from all three categories-curriculum, professional development and “other tools”. Those instructional resources that were made use of surfaced largely as face to face interactions in the form of professional development and other tools. This may be in keeping with the way the two teachers interacted on a regular basis. Their units were often co-planned, and they shared early in the study that a significant amount of their planning happens in the form of quick back and forth check-ins, text messages, and weekend calls.

4.2.1 Curriculum as an instructional resource

The curriculum was the least frequently accessed of the instructional resources. In this study, the curriculum is represented by the unit plan for the animal unit as it was taught two years ago when the teachers last taught 4th grade in the loop. For coding, I chose to consider the actual probe itself a part of the curriculum. Teachers identified and embedded both probes into the unit before being taught. While this study was partially the impetus to do so, the probes were not embedded just for this current iteration of the unit and will remain a part of the unit for future years.

In the first planning session, two prime examples of the curriculum as an instructional resource emerged in a single section of the discussion. In this segment of discourse, after talking about nutrition and the types of food the Divos eat, the teachers are grappling with questions about the actual structure of the Divos' mouths in the *Habitat Change*.

> **Teacher A:** Wait, the divo doesn't have teeth.
> **Teacher B:** Oh, it doesn't…
> **Teacher A:** Does it say that?
> **Teacher A:** It doesn't say it. It doesn't say that. You're right. So I'm just assuming.
> **Teacher B:** Assuming? But it might not.
> **Teacher B:** [reading probe] A small, short furred, grey animal called the divo. And the habitat is warm and provides plenty of divo food, tree ants. They live high in the treetops hidden from predators.
> **Teacher A:** Yeah, it doesn't say that. I guess I looked at the picture.
> **Teacher B:** But it could be-- I would say, in general, it's a misconception from a lot of kids that teeth and mouthpieces-- is that how you say it? Or mouth, mouths, whatever?
> **Teacher A:** Anatomy of the mouth.
> **Teacher B:** Teeth and the anatomy of the mouth can adapt quickly. But really these are adaptations that happen. And that's what the bird thing will address, right? The adaptation of the beaks?
> **Teacher A:** Very slowly. The difference in beaks?
> **Teacher B:** So on this one-- yeah. Bird/beak activity can address this misconception.

At the center of this discussion was the question of whether the Divo had the appropriate teeth for eating seeds. Teacher A believed that the Divo did indeed have teeth and needed to return to the probe in the unit plan to make use of this instructional resource. She read through both the narrative, the options, and the image to confirm that she was incorrect in her thinking. Teacher B vocalized that students appeared to have a misconception around the idea that the anatomy of the mouth can adapt quickly. The teachers subsequently identified and made use of *the Bird Beak Simulation*, an already established lesson from the animal unit, that would address this misconception because of its focus on the adaptation of the beak to the bird's food.

4.2.2 Professional development as an instructional resource

The second instructional resource category is professional development. Three unique professional development resources emerged as a part of the coding. The first professional development resource was the K-5 Science Specialist. The specialist was regularly available to the teachers as a resource throughout the unit and was present during the planning sessions as a resource for the teachers should they choose to make use of her. The second professional development resource was the supplemental teacher materials for the probes. These materials were the focus of professional development that took place over the preceding summer for all elementary teachers at the school. Knight-Bardesley and McNeill adapted the original Design Capacity Enactment framework to incorporate similar summer professional development that preceded (Knight-Bardsley & McNeill, 2016). The last professional development resource was myself, the researcher. Despite best attempts to remain just an observer during the planning sessions, the teachers did turn to me as a resource. The teachers using me as a resource occurred in connection to the facilitator role I played during the summer professional development. In this capacity, I served as an assessment specialist.

Instructional resources surfaced in Planning session 1 in two distinct ways: professional development and curriculum. The science specialist served as the sole professional development resource during this session. During two distinct clusters of turns in the discussion, she provided specialized content knowledge as a response to questions asked by Teacher B. These exchanges occurred during the planning session while the teachers were considering the misconceptions found in student reasoning from the first *Habitat Chang*e probe. Some students believed that the Divo animal in the probe could suddenly switch to eating seeds after all the ants disappeared, their current source of food. The excerpt below picks up after Teacher B recalled an article she read about how the pancreas of Homo Sapiens was larger than Neanderthal's because of the increased amount of grains in our diet as we transitioned from hunter to hunter-gatherer.

> **Teacher B:** …So animals' organs are adapted, are specially-- is that true? Specially adapted to their diet and change to diet happens slowly also. Is that true? Or can the animals switch, or does it depend on the animal?
> **Teacher A:** Depends like the Divos how adaptive they really are. Historically, the animals that are extinct, many of them were because they couldn't adapt to food and--
> **Teacher B:** To changes.
> **Teacher A:** --habitat changes, right?
> **Teacher B:** Quickly. So that's a slow thing too. Diet changing is something that happens slowly. Not quickly. Does it depend on if you're carnivorous or omnivorous? Is that something to look up? Carnivores, herbivores.
> **Specialist:** Yes. If they're used to looking-- if they're used to hunting meat, and then all of a sudden that source of meat is gone, they're going to have to try new foods. So, will they know to try new foods? To try plant-based foods versus--
> **Teacher A:** Or a different type of animal food.
> **Specialist:** Right.
> **Teacher B:** So, they could quickly-- a carnivore could probably more quickly adapt to eating a different type of meat than a carnivore could adapt to being an omnivore and eating plants. Do you think that's true? Meat is muscle tissue. It's pretty much the same with all mammals, right?
> **Teacher A**: That's a great question.
> **Teacher B:** Well, you know--

Specialist: With the amount of energy that a carnivore gets just from the meat versus the amount of energy that you get from a plant, the biomass has to be a lot more to equal the energy of the meat product.

The excerpt highlights two essential exchanges that both began with a question that demonstrated Teacher B making use of instructional resources. In the first turn, Teacher B inquired about the speed of diet change for animals. The exchange suggests that Teacher B's personal resources were insufficient to answer the question herself, particularly subject matter knowledge. Teacher A steps in and responds with an example of extinct animals and their inability to change quickly This demonstrated an opportunity where the subject matter knowledge of a singular teacher was insufficient to answer a question, yet the collaborative nature of the planning session allowed for collective subject matter knowledge to address the question at hand. This happens again when Teacher B asks about carnivores and omnivores in the fifth turn. However, this time, Teacher A does not interject with her knowledge. Instead, the question is left unanswered, and the specialist responds to the question posed.

Another example of the teachers making use of professional development as an instructional resource occurred at the beginning of planning session two as a review of the new Amphibian probe. The teachers decided that they did not like the original *Is It an Amphibian*? probe and decided to modify it and create their own. Despite having done this, they turned to both the science specialist and myself, the researcher, for significant input about all aspects of the probe. I highlight their initial exchange with me below.

Teacher B: So, we thought that we would probe it this way. What do you think? Did you get a chance to look at it? So, pictures with--
Researcher: Yeah, I think it's good.
Teacher B: Different characteristics. So, this one is at a stage. Do you think it's too focused or do you think it's--?

Researcher: No, I think it's good. So, the one thing I'd say is … those distractors are reptiles. So, I would think maybe to throw in some-- we had talked a little bit maybe put a worm in, or a slug, a lizard or other things like turtles that have soft, wet qualities like-
Teacher A: Moist?
Researcher: Yeah, moist. Yeah, moist animals.

In this excerpt, we see Teacher B concerned about the focus of the modified probe. As described in Section 4.1, the teachers narrowed the focus of the probe significantly. Their modified version only included amphibians at various stages of their life cycle and some other reptiles such as snakes. The feedback provided by the researcher was consistent with the research recommendations of the probe's supplemental materials. I felt it was essential to validate the teachers' efforts to modify the probe and make it more relevant to their focus. However, I also felt that there were some missed opportunities to use the related research found in the supplemental materials to support their modification efforts. Consequently, my feedback as a professional development resource provided a suggestion for expanding the narrow focus of the list in ways that were consistent with their focus. My feedback also mirrored that of the related research which highlighted studies on the common misconception among children that lizards are amphibians (Allen, 2010) and that students commonly misclassify turtles as amphibians(Yen, Yao, & Chiu, 2004).

Despite the previously described missed opportunities to use the probes' supplemental materials, in the second round of individual interviews each teacher identified the supplemental materials as a resource. When asked directly about their use of the supplemental materials, Teacher B briefly shared "We did on the Divo one [referring to the *Habitat Change* probe], but I did not on this [referring to the *Is It an Amphibian?* Probe]." Teacher B then went on to elaborate:

> I did read it ahead of time, and one thing that did inform my planning, in my planning with the interns is they stressed not to say, "Life cycle", all the time. That it's appropriate to say life cycle, like an individual organism has a life cycle, or has a ... no, no, has a life

Comparatively, Teacher A talked about the *Is It an Amphibian?* probe adding that:

> [referring to the supplemental materials] there definitely is helpful stuff on there and we shared that with the student teachers too as they were preparing to teach this lesson.

The two excerpts from the individual interviews demonstrate how the teachers made use of the supplemental materials as a pre-read and as a reference to pass along as a resource for the student teachers. One could characterize the teachers' engagement with these materials as one of convenience. The surface level context gleaned from early reading of the materials provided sufficient knowledge for the successful navigation of the probes implementation.

4.2.3 Other tool as an instructional resource

The final resource that emerged was the parent speaker. This resource was categorized under "other tools" as it did not fit neatly under the definition of the curriculum or professional development categories. I coded discussion turns involving the science specialist about the parent as professional development. After the parent speaker came to the class, the teachers used the presentation as a resource while they made sense of student data about hibernation misconceptions.

The end of planning session 2 focused on student misconceptions that still mostly existed after the administration of the second *Habitat Change* probe. One misconception that stood out centered around the students' ideas and reasoning about hibernation. This discussion quickly connected with the parent speaker who was a mammologist.

> **Teacher A:** And that's not all mammals. So, yeah, we have kids that said mammals hibernate. It's just a mammal thing.

Teacher B: And there's many mammals that do not. I feel like that's the advantage of most mammals is that--
Teacher A: They don't have to.
Teacher B: --they don't have to because they're resistant to-- and that's what the mammologist, talked about how mammals evolved from-- was it from reptiles was he saying? Do you remember that part? In the beginning, he was saying their big advantage was that they didn't have--
Teacher A: I remember that slide, but I can't remember if it was-- it must have been reptiles.
Teacher B: I took notes. He was talking about-- was it that they didn't have to--?
Teacher A: They looked like a dinosaur, right, in that picture?
Teacher B: They didn't have to--
Teacher B: This is what he was talking about, as far as how they adapted mammals coming from a branch of reptiles. And he said, [reading notes] "Early mammals had a big jaw, that moved back." Oh yeah. Instead of going up and down it moved--
Teacher A: Backwards.
Teacher B: --backwards.
Teacher A: I think that's what he said.
Teacher B: And then they had specialized teeth, like whereas reptiles have all the same teeth. And so mammals starting evolving these different types of teeth so that they could eat a wider variety of food, and then he also said that mammals are the only ones with three little ear bones. So mammals are really good at hearing, compared to other vertebrates. And that was the advantage, right, of mammals [inaudible] because I thought you'd talk more about cold-bloodlessness and hibernation, but really he's saying the advantages were more that they could hear their prey, or they could hear when predators were coming.
Specialist: Exactly.
Teacher A: Yeah. Then he went into the bats.
Teacher B: Then we went to how they have teeth for chewing and grinding. Because he wanted to point up to them that you could tell what type of an animal is an omnivore, or carnivore, or herbivore, based on looking at the teeth, right? Incisors are for biting and cutting, molars are for chewing and grinding.

It is here that the other tool category of instructional resources surfaces. As illustrated in the discussion, the teachers began referencing that parent speaker’s presentation as a resource. Since this presentation was neither part of the original curriculum nor teacher created, it was largely inaccessible to the teachers except for limited recollections. The teachers eventually turned to notes they had from the presentation that served as a resource to move the conversation forward. This demonstrated how an unforeseen resource, the parent speaker mammologist, served as a content expert for the teachers and ultimately a new instructional resource. The teachers make use of this other tool resource in a similar manner to how they made use of the supplemental materials as a professional development resource. The engagement with the resource is characterized by its immediate accessibility.

4.3 TEACHER RESOURCES

In this section I address my third inquiry question: "What teacher resources—pedagogical content knowledge, subject matter knowledge, and beliefs—do teachers express and how do they make use of these resources in response to data collected from embedded formative assessment probes?" Teacher resources are broadly defined as the resources that are internal, or personal to the teacher, and are divided into the three categories: pedagogical….. Analysis of coding revealed that the case study teachers expressed teacher resources from all three categories-pedagogical content knowledge, subject matter knowledge, and beliefs. Pedagogical content knowledge accounted for 68.7% of all the discussion turns coded for containing an expressed teacher resource. Subject matter knowledge and beliefs accounted for 15.9% and 15.4% respectively. Evidence of the case study teachers making use of each teacher resource in response to data collected from the probes surfaced in collaborative planning sessions and in both teachers' individual interviews.

4.3.1 Pedagogical content knowledge as a teacher resource

In this study, pedagogical content knowledge (PCK) was viewed as being composed of three sub knowledges: content and curriculum, content and students, and content and teaching (Ball et al., 2008). The PCK coded in discussion turns ranged from simple examples to complex lines of thought that wove PCK together with other resources. Section 4.3.1.1 highlights multiple examples of how the teachers expressed PCK through discussion of the misconceptions that arose from the probe around the concept of hibernation.

4.3.1.1 Hibernation-related PCK discussions

The PCK examples from the three sub-categories were expressed throughout all five planning sessions. Knowledge of content and students (KCS) represents an aspect of PCK that includes misconceptions, student strategies, and student understandings of content. In planning session 1 the teachers engaged in a brief discussion about hibernation for the first time. In the example we see the teachers recognize that a large number of students think that the Divo from the probe can just start hibernating. They also hone in on the student response that “most” animals can hibernate. Consequently, the teachers use this data to begin making a list of misconception they must tackle, including what animals hibernate and how quickly an animal could start hibernating if it never has before.

> **Teacher A:** Two people said their hair grew longer.
> **Teacher B:** [reading a student’s response] "Like most animals, they will hibernate."
> **Teacher A:** Most animals [laughter]?
> **Teacher B:** Something [inaudible]. "Then when it got warmer, they stopped hibernating." So we need to talk about hibernation. That's clear. Because I don't think--
> **Teacher A:** And adaptations. Most of mine said they died or hibernated.
> **Teacher B:** So misconceptions, [teacher b reads aloud notes she is typing] most animals hibernate. Hibernate. Most animals hibernate. Only some do. And then adaptations occur within an animal's lifetime.

> That's a misconception because that happens slowly over many generations. What else would you say is a pattern? Well, I guess we should address all of them. Let's go with the most common ones first.

Although brief, this exchange established an early baseline for what misconceptions the teachers would need to be cognizant of moving forward. In Planning session 2 that immediately followed the second implementation of this probe, we continue to see KCS surfaced through the careful examination of student responses:

> **Teacher A:** [reading from student responses] So here's a misconception: "The Divos hibernated through the cold period or they died." Those were his responses. And his response to them hibernating is because they are mammals. It sounds like he thinks all mammals can hibernate. And then, "Divos died because they're fake, made-up animals [laughter]." That was his second response.
>
> **Teacher B:** Oh, here's someone that explains that, "They've probably switched to seeds because they're almost like ants, small and skinny-ish." So she's assuming it's the same size. "I also think that they would start hibernating because they might not be able to live in the cold. So they can dig holes and hibernate, and come out in the spring."
>
> **Teacher A:** I have another response like that. So she said either switch or hibernate, and I said there were-- it said that there were plenty of seeds on the ground, so the Divos probably could have switched to eating seeds. But if the Divos could not switch to eating, they could hibernate until it gets warm again.
>
> **Teacher B:** This one shows no change in thinking.
>
> **Teacher A:** I have another one. "Because most animals hibernate in the cold."

In addition to this, the two teachers also expressed and made use of their knowledge of content and teaching (KCT). KCT is the knowledge of how to teach. In the planning session we see it emerge as teachers consider what to do with the knowledge from KCS. We see the teachers grapple with the fact that numerous students still think that hibernation is still something that the fictitious Divo can spontaneously decide to do. Teacher A is the first one to note that there is something “funny” about how the students responded. In this exchange Teacher B had a lot of ideas about their early noticings of student responses in the second implementation of this probe. In particular, she raises concerns about what was both done to address the hibernation misconception and what was not done:

> **Teacher A:** So it's funny. So we gave them the exact same thing. But we wrote at the top, [reading from the probe] "second response, take another look at the divos. Your response should reflect how much you've learned about animals from our studying." So it's funny. Because I'm hearing a lot of the words that we talked about like hibernate. We talked about adaptations, and the kids-- a couple have mentioned, "They'll probably hibernate."
>
> **Teacher B:** And I noticed that kids-- a lot of my-- I was interested that a lot of the kids were like, "Well, the animals are going to die." And now, more of the kids selected multiple things and can explain, like this one just says-- because they had originally just said that the Divos died.
>
> Although, I am concerned that they didn't-- he didn't then say the Divos died. So does he think that everything now can adapt? Do you actually create a new misconception? That before the animal didn't have the conditions that were right, is it possible that now, they think everything can adapt? Like everything has at least a chance of adapting?
>
> But it just says, "The divos grew longer fur because they couldn't start hibernating during the winter." So he said, "I acknowledging that you can't just start hibernating." And then he said, "But they can just start growing their fur longer [laughter]," which is kind of funny, isn't it [laughter]. Yeah. So I have a lot that have multiple answers and some of them don't say that they die. But a lot of them are saying that, "Or they'll die [laughter]," and hibernation seems to be a popular one. So maybe we need to spend more time talking about what hibernation is. You can't just make that decision obviously to buckle down for the winter.
>
> **Teacher A:** You have to understand the process and the preparation that goes into it.
>
> **Teacher B:** Oh, so you know what we didn't talk about is, we didn't talk about behavioral adaptation versus physical. So I bet that's why. Hibernation is a behavioral adaptation, and in the past, we talked about that.
>
> **Teacher A:** [agreeing about last time] We did talk about that.
>
> **Teacher B:** Behavioral versus physical adaptations. So maybe we need to talk about-- because that book, that little book actually addresses both. But I bet the kids were focused more on the physical because of the activities [referring to the bird beak simulation, the bird skull lesson, and the owl pellet activity]

From this exchange we see that Teacher A drew upon her KCT and expressed the notion that the misconception that everything can adapt may have been inadvertently created because of the teachers focus on adaptations. Similarly, Teacher B identifies that they only discussed adaptations broadly. She hypothesizes that since they did not explicitly discuss the differences between physical and behavioral adaptations that students learned less about behavioral adaptions since the activities focused more heavily on physical adaptations.

As planning session 2 progressed the discussion later returned to the topic of hibernation. However, this time the nature of the PCK expressed centered more around knowledge of content and curriculum (KCC). KCC is knowledge of the available materials teachers can use to support student learning.

> **Teacher B:** So that's the value of the probe, is to catch the stuff that you overlook, that you knew was important, but you didn't focus on. As an adult, you're aware of all this stuff, but the kids, you don't realize, I think, always the implications for the activities you choose, right? By choosing only to look it physically-- but I guess because of hype it would've been harder to show them behavioral adaptations, we would've had to get a series of video clips instead of like the bird skulls from the museum or the bees which is somewhat--
>
> **Teacher A:** Somewhat physical.
>
> **Teacher B:** --with their hand out. Although there is probably a way with the bird beaks simulation, you could talk about behavioral adaptations. Like, what if you were allowed to-- I don't know, change the way that your beak scooped up. Like the kids were all using their beaks the same way, like forward and backward, but some of them didn't notice you could get more if you scooped to the side. Or something like that. So there might be a way with something like that to talk about the behavioral adaptations, but things like hibernation happen over such a long period of time it's hard to show that to kids. And something concrete. So, it makes sense that the misconception shifted away from just any variety of misconceptions to the ones that are less concrete.
>
> **Teacher A:** Which are behavioral and--
>
> **Teacher B:** Yeah, behavioral. But the amphibians, where they're moving next, like there's a number of ways that we could-- like frogs hibernate, [asking group] right? Don't most amphibians hibernate in the winter? Don't they dig under the mud?

It is here that teachers discuss for the first time, how the concept of adaptations may need to continue into the next section of the unit. Their awareness of the affordances of the upcoming amphibian section of the unit allowed for them to come up with potential ways to continue challenging student misconceptions on hibernation.

4.3.2 Subject matter knowledge as a teacher resource

In this study, subject matter knowledge is viewed through the framework set forth by Deborah Ball. While there are examples of common content knowledge and horizon content knowledge, this study found specialized content knowledge to be the primary resource that teachers expressed and made use of. Common content knowledge is defined as knowledge necessary for everyday such as fire is hot, or animals are alive. As suggested by its name, specialized content knowledge is knowledge that requires a teacher to have a deeper understanding then simple common knowledge. While teachers drew up on subject matter knowledge in all planning sessions 1 and 4 contained particularly significant clusters of subject matter knowledge expressed by the teachers. Each of these planning sessions aligned to a different probe. Planning session two also contained some subject matter knowledge however the amount was minimal as the majority of surface teacher resources were various expressions of pedagogical content knowledge. All 3 of these planning sessions marked the beginning of the observed formative assessment sections. Each of these sections describe in section 4.1 represented a common theme in the way that teachers responded to the data from the formative assessment probes. Consequently, the way in which subject matter knowledge was expressed in each of these sessions is representative of the overall theme of the formative assessment section. For example, the small amount of subject matter knowledge in planning session 2 is directly related to the large amount of time spent by the teachers making sense of misconceptions, expressing pedagogical content knowledge.

Planning session one immediately followed the first implementation of the *Habitat Change* probe. This session was the first opportunity that the teachers had to work with real student data from the probe. Since the teachers did not modify the habitat change probe they worked with the probe as written by the author. The fictitious nature of the Divo animal in the probe caused the teachers to occasionally surface their own subject matter knowledge as they made sense of student responses. This is seen in some of the early exchanges as the teachers struggle with the question of whether the Divo is a mammal. While the probe does not explicitly say this, it is implied because the animal is covered in fur:

> **Teacher B**: So, if mammals care for their young, then they probably would-- if it's a mammal, then it could adapt, right? Do mammals always take care of their young?
> **Teacher A**: That's one of the--
> **Teacher B:** I think that's one of the trademarks for them, right?
> **Teacher A:** Does the platypus take care of theirs, and [inaudible].
> **Teacher B**: I don't know.
> **Teacher A:** ...if they're born from eggs?
> **Teacher B:** That's a good question.

As we see a number of the examples of subject matter knowledge being expressed were done so through statements with inflection at the end indicating that wow the teachers had some knowledge in this area they may have not been the experts in it that was required to navigate the student misconceptions. In fact, some of the subject matter knowledge expressed demonstrated some of the misconceptions described in the supplemental materials of the probe. We saw in section 4.2 that the teachers' use of the supplemental materials were limited to an initial reading and were not gone back to end used as a reference during the planning sessions. The supplemental materials summarize research findings from Rosalind Driver et al. that "Students appear to confuse an individual’s adaptation during its lifetime and inherited changes in a population over time. A large number of students appear to adopt a Lamarckian view of adaptation” (1994). Teacher B expresses this very misconception of Lamarckian thinking where one believes that learned adaptations can be passed on to the next generation. The discussion in the planning session centers around the notion of whether or not the Divo can learn to build a nest or burrow for survival:

> “So, they don't quickly adapt. This [referring to the ability to figure out how to build a nest or burrow] may only apply to one or two divos and then they would have to develop a way to pass the knowledge on to their offspring, right? Because what if one or two figures out to live in the leaves? That would be likely. But then they would also have to-- we don't know enough about their-- sophistication or how they're organized.”

Planning session 4 immediately followed the final implementation of the *Is It an Amphibian?* probe. While the other probe was used as written by the author, this probe was adapted by the teachers to best fit the content they were teaching. The teachers were more familiar with the content of the probe and the reasoning for the inclusion of each of the items. As a result, the teachers surfaced little to no uncertainties about the actual subject matter knowledge associated with the items of the probe. Consequently, as evident from the opening discussion turn of planning session 5, Teacher B came to the session with a list of disciplinary content ideas all ready for discussion. This is in keeping with the theme identified in section 4.1.3 “reviewing and assessing.” In their on the fly development of a task for the purposes of reviewing the teachers expressed subject matter knowledge through the development of the distinguishing categories of vertebrates some of which included: type of birth, warm or cold blooded, body covering, and how they feed their young.

4.3.3 Goals and beliefs as a teacher resource

In this study goals and beliefs are captured in a number of different ways including surfacing in the planning sessions, elicited in the individual interviews, and evaluated through the completion of the teacher belief inventory in the final interview. I first discuss teacher beliefs through the broad lens of the teacher belief inventory. The inventory was completed during the interview at the end of the unit. The inventory provides a profile of each teacher ranking their responses along a continuum from traditional to reform-oriented teaching. While the inventories don't provide a singular label about the teachers' belief, it does provide the big picture of how the teachers think about teaching students' science.

Table 10. Teacher Belief Inventory for Teacher A

Question Topic	Traditional	Instructive	Transitional	Responsive	Reform based
maximize student learning	routine for when we're starting an activity map out chunks of time…so that we get through everything		honor the kid's need or want to stay on a topic longer		
role as a teacher			them enjoying it and liking it is a big part of the hook, getting them interested	introduce things to the kids, listen to their questions and their comments, try to bring that into what I have planned	
when your students understand		written reflections are one big then either through discussion or checking of the notebooks			
what to teach and what not to teach?				what the kids show us, or where their interests are, a lot is based on the kids use the standards for the level.	

Table 10 continued

when to move on to a new topic				Kids show interest, or they are still engaged, I have a hard time moving on getting tired of something, that's also an indicator	
learn science best		follow a routine through our discussions, pointing out things that I want to make sure they know and understand.	being active and involved, engaged in the lessons		
when learning is occurring			reflections engagement to a degree		

The teacher belief inventory shown in Table 10. showed that Teacher A held several beliefs that ranged from traditional to transitional to responsive yet did not include any responses that were categorized as reform-based. While a few of Teacher A's beliefs were either responsive or transitional there were still a number of responses that favored the more traditional end of the spectrum. It is clear though from the inventory that despite responses across the inventory spectrum, Teacher A often has students in mind when making decisions about instruction.

In addition to the inventory the teacher also expressed a number of beliefs in the individual interviews. One such beliefs centered around the teacher’s thoughts about the composition of her class and their knowledge from previous years’ instruction where she surfaced beliefs more aligned with the instructive category from the beliefs inventory:

> “this is a very chatty group, a lot of behavioral issues. I'm not sure how much they learn. I know they did insects last year and they've done animals in the past. I wasn't sure how much. It seems to be spotty.”

We also get a glimpse into the teacher’s belief about how her students share what they have learned. The teacher emphasizes that students may often struggle to fully explain or demonstrate their knowledge through writing and that her perception of their knowledge is strongly informed through their discussions and mirrors her transitional beliefs from the inventory about when learning is occurring:

> “I think that we've cleared up misconceptions just because we've talked about it so much and I'm hearing a lot of what the kids are saying. I think sometimes when you leave it up to them to just write, some of my kids aren't strong writers, or they're resistant to writing and I think that holds back on what they share and I think that they do know more, if it's not written, I don't believe that that means they don't know it.”

Table 11. Teacher Belief Inventory for Teacher B

Question Topic	Traditional	Instructive	Transitional	Responsive	Reform-based
maximize student learning		maximize the amount that the kids will absorb and retain better to scale back or do less ambitious and then have the kids learn it deeply.			
role as a teacher					to curate experiences…kids to discovering things on their own. An expert, but more as a shepherd
when your students understand		you have to have a one-on-one conversation	during whole class discussions, some stuff will come out.		
what to teach and what not to teach			what I've done in the past if I know something went over really well in the past		
when to move on to a new topic				they were really engrossed, I perceive them to be learning a lot	
learn science best			most of them will do like activities, hands-on activities.	they have to try to articulate it, like they've seen some pattern. giving them new data, asking them some hard questions is when they solidify their understanding.	
when learning is occurring		completed the task	through discussion or when the kids are saying they're done with something, but they really are not		

Teacher B presents differently along the teacher belief inventory. Her responses range from instructive on one end to reform based on the other. The majority of her responses are around the center of the continuum and could be categorized as transitional. Of particular note, is how Teacher B perceives her own role. As indicated in the second row of the table the teacher expresses her role as three-fold: curator, expert, and shepherd.

Another item of interest in the responses is the teacher's indication that she knows when her students understand by having a one on one conversation. This is consistent with the way in which both probes were adapted asking students to explain their thinking item by item, much in the same way one would engage students in one on one discussions.

Teacher B also expressed some strong beliefs about her expectations of where she thought students should be entering the intermediate grades. The teacher recognizes that students in primary grades were engaged in good science process skills such as observation. However, her focus on mastering definitions reflects very traditionally aligned beliefs:

> "Yeah. I didn't think that I would have had to do that [referring to spending time on the differences between amphibians and reptiles], because we already had covered it in a brief survey in the beginning, and the students have an animal unit in primary. I would have expected that they would have figured out the definitions in primary, but I think they just get a rudimentary understanding of animals. I think they're just making observations, and then maybe talking about those things, but it's not mastered."

This is consistent with the amphibian topic within the animal unit. The concept of classification is addressed, but it is done simply through the tasks of differentiating between amphibians and reptiles. The teacher goes on to share her surprise about a particular student's response to an iteration of one of the probes. In doing so she expresses her beliefs that certain students will perform certain ways:

> "So one student that I was really not expecting to be successful in this gave some of the most insightful observations. So I think that surprised me. And that I think is a really good thing about something that's short, but very targeted."

We also see teacher B express some strong beliefs about the impact of privilege and social status of the student population and its direct relationship to her beliefs about their performance on the probes. While the picture painted may be an accurate representation of students' activities outside of the classroom it also represents the teacher's perception that students are of a homogeneous nature:

> "I also feel like our kids are privileged. They go to the zoo, they read books with their parents. They watch science shows. If they're getting screen time, they're not just watching junky programming. They're usually watching a NOVA special. So, yeah, I was a little surprised that they didn't know consistently."

Overall, we see the teachers' beliefs fall across the full range of the inventory. This is also true with beliefs expressed by teachers through interviews and planning sessions. However, their beliefs tend to center around the transitional category suggesting a "focus on teacher/student relationships, subjective decisions, or affective response" and an overall guided inquiry approach to science.

4.4 CRAFTING NEXT STEPS

In this final section, I address my overarching inquiry question, "How do teachers make use of resources available to them to interpret student responses from embedded formative assessment probes and craft next steps in instruction?" However, before I do talk about the tasks that best help me answer my overarching question I first want to address why I am not discussing all tasks. Of the tasks that were the subject of this case study, only those related to the data from the formative assessment probes were examined for inclusion in this section. Fifteen lessons were coded and analyzed. The final two tasks of the unit were referenced significantly more than the other tasks and are highlighted below.

A few of the improvised tasks that were named for the concepts they cover can be traced back through the various planning session discussions. However, the discussion of these concepts in earlier planning sessions is not directly linkable to the tasks. A prime example of this would be the *Physical vs. Behavioral Adaptations Venn Diagram* which took place toward the end of the unit. This task occurred during the “changing everything” period of the animal unit when the teachers were attending to misconceptions from two different probes. The concept of physical versus behavioral adaptations can be traced back to discussions from the very first planning session when teachers were only aware of misconceptions from the *Habitat Change* probe. While one could infer connection and causation between the various discussions of physical versus behavioral adaptations, the fact remains that there are insufficient surfaced connections between all of these references leading up to the task. This highlights how improvised tasks may be talked about very little proceeding their development and implementation.

During planning session 4 the *FRAMB Graphic Organizer* was developed on the fly in direct response to the persisting misconceptions that arose after the second implementation of *Is It an Amphibian?* probe. Examples of the misconceptions that surface early in planning session 4 focus on the concepts of vertebrates versus invertebrates and the distinguishing characteristics between amphibians and other animals like reptiles and worms. The identification of these persistent misconceptions led to a decision by teacher A to review the characteristics of amphibians. What followed was an intensive back and forth dialogue where the teachers drew on various resources, making use of them to craft an improvised next step in instruction. A summary of those resources contributing to the development of the graphic organizer are found in Table 12.

Table 12. Summary of resources teachers made use of for *FRAMB Graphic Organizer*

Pedagogical Content Knowledge	• Identified persistent misconceptions; vertebrate/invertebrate, life cycle, protection, etc. • Identified different ways to organize information • Identified different tasks to review the characteristics
Subject Matter Knowledge	• Identified categories for distinguishing characteristics between vertebrates
Goals & Beliefs	• "The rich got richer, and the poor stayed where they were at." • "the kids that already knew a lot were more excited and engaged in the activities." • Need to continue using graphic organizers; fourth graders need more practice with them • It is hard to undo misconceptions. Much easier to start fresh
Curriculum	None
Professional Development	• Made use of science specialist during brainstorm of FRAMB structure.
Other Tools	None

In contrast to the chance capture of the improvised FRAMB Graphic Organizer, adapted and offloaded tasks potentially provide more opportunity to observe the teachers discussing their application and the ramifications of the data collected from the planning sessions on the implementation of such tasks. The final *Create Your Own Amphibian* project provides a prime example of how an adapted task from the previous unit became the focus of numerous lines of discussion.

Table 13 provides a summary of the resources surfaced throughout the study. Discussions leading up to this adapted task converged from 3 original lines of discussion. The first occurs during planning session 3 which occurred after the initial implementation of the amphibian probe. This planning session occurred at the start of the changing everything formative assessment section identified in 4.1.2. It was known at the start of this amphibian related topic within the science unit that the final projects would be related. Hence all tasks and consequent improvisations within this section were done with the goal of preparing students for this final task.

Table 13. Summary of resources teachers made use of for *Create Your Own Amphibian*

Pedagogical Content Knowledge	• Identified misconceptions surrounding the behavioral and physical adaptations • Explicitly incorporated adaptations into aspect of the task. • Reviewed amphibian characteristics through numerous tasks
Subject Matter Knowledge	• Knowledge of biomes, adaptations, and amphibians.
Goals & Beliefs	• Expressed a belief that the intersection of the science and social studies provided a strong reason to repeat structure next time.
Curriculum	• Lesson plan for the task from the previous time the unit was taught.

Table 13 continued

Professional Development	• Science Specialist was utilized as a content resource.
Other Tools	None

Additionally, the ongoing focus on the topic of adaptations surfaces itself in the final project. Teachers indicated that the project was adapted to purposely incorporates adaptations. Although not a focus of this study, the social studies unit on biomes and landforms also became a part of this final project setting the stage for the placement of the adapted amphibian.

4.5 CONCLUSIONS

A number of big ideas emerged in the findings of the study. Of foremost significance was the varied ways that teachers responded to the data from the embedded formative assessment probes. We saw teachers respond initially with little change to their planned instruction. Teachers surfaced beliefs that existing instruction would address the misconceptions related to adaptations. When presented with further data after the second administration of the probe later in the unit teachers began to shift he way they approached planning their instruction. The persistence of the same misconceptions revealed to teachers that student thinking on certain concepts did not change or shift after the planned instruction. Consequently, the teachers entered a phase where significant changes were made. This period was typified by the improvisation of new lessons and tasks and the use of no prior lessons exactly as taught the previous year.

Of other significance were the resources upon which teachers drew and made use of. While we did see teacher make use of instructional resources from time to time, It was clear from the discourse, interviews and unit artifacts that teacher relied heavily on their own personal teacher resources. The most prevalent of these was the teacher's own pedagogical content knowledge. The improvisation of new lessons and the student data revealing misconceptions required teachers to make use of the their PCK in light of the lack of affordances provided by the existing curriculum. Teachers also surfaced a number of strong beliefs and goals about students, abilities, and instruction. Regardless of their accuracy, these beliefs provide revealed the strong influence and opinions that teachers brought with them to the table.

5.0 DISCUSSION

This case study provided numerous opportunities to examine these master level teachers as they engaged in the new practice of embedding formative assessment probes into their instruction. The discussion of the findings is not generalizable due to the small size of the case study. Instead, the study provides insight into the ways that teachers make changes to instruction and the resources they draw upon in doing so.

Numerous themes emerged as they progressed through the unit. One such theme was that of change. The animal unit underwent significant changes with more than 85% of the lessons new to the unit or adapted from old lessons. Nearly half of these changes were not related to data collected from the formative assessment probes suggesting that the animal unit was undergoing significant changes even without the added data from the probes. Brown (2009) defines PDC as a teacher’s ability to perceive the affordances of the curriculum. Given this, one might conclude that the teachers perceived little affordances from the pre-existing curriculum which was structured according to the five classes of vertebrates: fish, reptiles, amphibians, mammals, and birds. Remnants of this old organization can still be found in the animal unit with the heavy focus on amphibians in the latter half of the unit. The earlier half of the unit that focused on adaptations incorporated both birds and mammals. Given that the unit appears to be in a transitional state and may very well likely undergo further significant changes in future years.

The study findings aligned and reinforced numerous areas of the literature. One area where connections with the literature emerged was during the adaptation and improvisation of instructional sequences in the form of teachers drawing upon resources that were familiar to them or already existed elsewhere. This is consistent with what Falk (2012) found when examining teachers' use of formative assessment and next steps in instruction. Although not related directly to the data from the probe, we saw the teachers incorporate new lessons into the unit that drew on the wildlife guide resource that was tucked away on a shelf. Similar occurrences surfaced when activities and lessons were discussed from old teaching partners. These familiar resources were drawn upon due to their content and alignment to the science concepts taught. Teachers also drew upon familiar resources in the way of strategies. Throughout the study, both teachers repeatedly utilized graphic organizers and discussions both of which appeared to be a core strategy from the teachers' toolboxes.

In one of the stronger studies focusing on next steps in the instruction, Cowie & Bell (1999) identified three ways that science teachers acted on the results of their embedded formative assessments : science-referenced, student-referenced, and care-referenced. As described below, these categories provide a compelling means to consider the ways that the case study teachers responded to the evidence elicited from the formative assessment probes. Additionally, the ways that teachers acted on the formative assessment reinforced findings from Cowie & Bell.

Science-referenced next steps in instruction centers around the idea that once misconceptions have been elicited by the teacher the response to this knowledge is to plan activities directly aligned to the misconception. A prime example of this occurred in planning session 2 when the teachers realized that their focus on adaptations in the *Bird Beak Simulation* was limited to physical adaptations. They concluded that the focus on physical adaptations in the simulation was the reason student corrected their misconceptions about physical adaptations only. Students thinking shifted away from physical adaption options and toward the remaining, unchallenged behavioral adaptation choices on the probe. Consequently, the teachers planned specific activities to address this misconception.

Care-referenced next steps in instruction seek to support some aspect of the social interactions between students in response to the information gathered from the assessment. We saw examples of this take place with the teachers as they surfaced their reflections on the student activities that engaged students in whole class discussions. While not evident from the names of many of the lessons, whole class discussion was a significant component. An example of this is seen in the *Frog vs Toad Venn Diagram* lesson. Students worked on completing the Venn diagram in small groups or pairs and then engaged in a whole class discussion. These were often talked about in a positive way by the teachers and described as being "rich" in nature. The consistent presence of whole class discussions in lessons related to the formative assessment probes suggests that this form of social interaction was valued and seen as having a positive impact on learning.

Student-referenced next steps in instruction centers on the teacher's awareness of the individual student's changing understanding and providing an appropriate instructional activity in response to this that reflects the next concepts the student is capable of learning. One might expect to see this in the form of differentiated instruction. The multiple choice list of options that students could select from in the *Habitat Change* probe provided ample evidence to support teachers in crafting differentiated instruction. However, we did not see any lesson or activity that was not "whole class" in nature.

My case study also reinforced findings of the report from the Center for Research on Evaluation, Standards, and Student Testing (Heritage, 2007) that found that in order for teachers to make decisions about next steps in instruction, they must draw heavily on subject matter knowledge and pedagogical content knowledge. The teachers in my study drew primarily from both. The teachers surfaced pedagogical content knowledge in the planning sessions most often. This supported the improvisation of instruction that took place throughout much of the latter half of the Animal unit.

The literature also offered some significant critiques of formative assessment. In particular, Coffey et al. (2011) argued that too much attention is paid to instructional activities and not enough time is spent on what is occurring in the classroom and what the teachers notice and consider. In section 2.3.1 I contended that embedded formative assessment in science provides teachers with stronger tools aligned with learning goals and research on student thinking. These tools tax teacher subject matter much less than informal formative assessment that takes place during instruction which is the focus of Coffey's critique. This case study provides evidence to the contrary by demonstrating that embedded formative assessment provides an opportunity to focus instructional activities and what teachers notice and consider. Embedded formative assessment affords teachers the opportunity to consider data from formative assessment outside of the actual classroom instruction time and during planning periods.

Coffey also critiqued the lack of disciplinary substance in the formative assessment field. This case study engaged teachers with an embedded formative assessment tool that was strongly grounded in disciplinary content and research. Formative assessment itself was not a new practice for the case study teachers. However, the use of disciplinary-rich, research-based formative assessment was a new practice and consequently provided an appropriate lens for this case study.

5.1 LIMITATIONS

In reflecting on this study, I consider the limitations of the Design Capacity for Enactment framework in this setting. In particular, two significant issues surfaced during the study that I believe considerably hindered the efficacy of the framework: the teacher-created nature of the unit coupled with the high occurrence of improvising and freedoms afforded to teachers through their school setting.

The teacher-created nature of the unit raised some questions during the study. One such issue was the limited nature of what it meant to offload. Offloading can occur for numerous reasons including inadequate personal resources (“I don’t know enough about this topic”), perceived benefits of the curriculum (“the curriculum does a great job at teaching this”) and affordances of curriculum (“I need to focus my time on another portion of the curriculum”). Knight-Bardsley and McNeill (2017) found that offloading provided teachers with the opportunity to step outside of their comfort zone. Teaching in this manner provides opportunities for growth and new knowledge that informs the personal resources side of the framework that may help the teacher to improvise in the future more effectively. A prime example of this is a first-year teacher grappling with a subject area where they have little experience. The teacher offloads most, if not all, of the lessons and in the process of doing so improves upon their personal resources.

In my study where teachers created an animal unit, offloading represents a teacher not making changes to the lesson at hand and teaching it the exact way that they had designed it in previous years. In this instance, the offloaded lesson provides few new opportunities for growth. In this scenario, adapting and improvising takes on a similar altered meaning. These approaches to designing instruction require the teachers to weigh their alternatives to the curriculum against the curriculum they created.

Another limitation that impacted the efficacy of the framework were the freedoms afforded the teachers within their classroom spaces and ultimately their school setting as a whole. I originally intended to address this through teaching binders and attending planning periods:

> "[excerpted from study overview methodology] Following each session teachers will be asked to keep a binder with lesson plans for the original unit as planned. Any changes that are made in response to the data collected from the probe will be annotated in the margins or inserted as new lessons. Additionally, teachers will be asked to commit to dedicating one planning period per week for science planning."

The original request to the teachers to keep a binder and record changes was countered with a request to utilize Google Docs where the unit planning was already taking place. It was a more convenient option for the teachers and placed less of a burden on their time. I agreed that this would be a sufficient alternative, however, the resulting shared document, ultimately lacked the structure and consistency to be used in a way that provided the data needed.

Committing to a single science planning period also proved too difficult to manage as the two teachers regularly engaged in casual conversations, text messaging, and phone calls/emails on the weekend for planning purposes. Ultimately, these limitations reduced the insight into the teachers' decision-making process and the resources they made use of in making these decisions.

5.2 IMPLICATIONS FOR PRACTICE

A significant implications for practice is directly related to elementary science teaching within the school setting. One thing that was abundantly clear after spending a semester with these two teachers was their significant personal knowledge. This came in many different forms and often complemented one another. Teacher A often described the two teachers as balancing one another. Teacher B had lots of ideas, and Teacher A was good at narrowing down the choices. These teachers demonstrated a working relationship that was indicative of mutually beneficial synergy. They relied upon one another for help with planning, teaching, reflecting. While these were normal behaviors for them, the practice of embedding formative assessment in this way was entirely new to them.

The planning sessions that took place after formatively assessing students created a space for the teachers to talk about student thinking and its implications for instruction. In my opinion, this space was of great importance. It provided opportunities for teachers to respond to student thinking in a proactive way before the end of the unit. It also created a space for the science specialist to passively and actively engage with the teachers. Such spaces and opportunities may prove valuable in other grades and of equal value for the specialist who, in my opinion, was considerably underutilized. These opportunities to engage with formative assessment data should not be limited to the scope of the unit at hand. In fact, as Wiliam (2014) suggests, the formative assessment feedback cycle can occur over semesters or even years. With that in mind, the school may benefit from the development of “handoff opportunities”. In this schools’s looping format, students have the same teacher for 4^{th} and fifth grade. The creates a situation where a given teacher only teach a particular unit once every two years. It is taught by the other set of teachers in the interim years. The handoff opportunities provide both sets of teachers the chance to discuss lessons learned from surfaced student misconceptions revealed through formative assessment.

Another implication for practice centers around future professional development opportunities. This study was preceded by a summer workshop that focused on an introduction to how embedding formative assessment can be used to gather evidence about student thinking. One resource that was a focus of the summer PD was the supplemental materials that accompany each probe. This resource was found to be underutilized during the study. There was no evidence that the teachers in this study made use of sections on *related research* and *suggestions for instruction* in the supplemental materials. In particular, the supplemental materials contained guidance on the misconceptions directly related to the final project of the animal unit: the *create your own amphibian* task. Throughout the unit, the teachers struggled to change student misconceptions about how quickly an animal can adapt and that they can do so by choice. The supplemental materials identify that exact task (design your own animal) as a possible next step in instruction. However, that recommendation comes with a warning that it may "perpetuate the misconception that organisms intentionally adapt" (Keeley et al., 2005). Future PD should focus on leveraging new insights teachers have into their students' thinking and how resources like the supplemental materials can be used to support informed decision-making during the unit planning and reflecting process.

5.3 IMPLICATIONS FOR FUTURE RESEARCH

While the limitation presented significant barriers, they also provide direction for how to use Pedagogical Design Capacity as a tool for better understanding teacher affordances of instructional resources in their decision-making surrounding next steps in instruction. While not the focus of this study, an additional area of interest emerged over the course of this study: task analysis. Furtak et al. (2016) found in a study about the development of a Formative Assessment Design Cycle that after four years teachers showed improvement in all areas except for the quality of tasks. Although not measured explicitly in this study, I believe that the tasks that were adapted or improvised were of similar quality to those they replaced. This is supported by the repeated use of given lesson formats and types of tools such as specific graphic organizers or the use of whole class discussion.

A more in-depth study of task selection when planning next steps in instruction could prove to be an area that sheds further light on how teachers continue to struggle. Research in the task analysis area would also lend itself to the PDC framework and vice versa, supporting inquiries into the connections between resources teachers draw upon and the quality of the tasks they select.

APPENDIX A

CHECKLIST OF ELEMENTS FOR OBSERVATIONS

1. The physical setting:
    a. What is the physical environment like?
    b. What is the context?
    c. What kinds of settings is the space designed for?
    d. How is space allocated?
    e. What objects, resources, technologies are in the setting?
2. The participants:
    a. Describe who is in the scene, how many people, and their roles.
    b. What brings these people together?
    c. Who is allowed here?
    d. Who is not here that you would expect to be here?
    e. What are the relevant characteristics of the participants?
    f. What are the ways in which people in this setting organize themselves?
3. Activities and interactions:
    a. What is going on?

b. Is there a definable sequence of activities?

c. How do people interact with the activity and with one another?

d. How are people and activities connected?

e. What norms or rules structure the activities and interactions?

f. When did the activity begin?

g. How long does it last?

h. Is it a typical activity, or unusual?

4. Conversation:

a. What is the content conversations in this setting?

b. Who speaks to whom?

c. Who listens?

d. Quote directly, paraphrase, and summarize conversations.

5. Subtle factors:

a. Informal and unplanned activities

b. Symbolic and connotative meanings of words

c. Nonverbal communication such as dress and physical space

d. Unobtrusive measures such as physical clues

e. What does not happen

6. My own behavior:

a. How is my role as a participant observer affecting the scene I am observing?

b. What do I say and do?

c. What thoughts am I having about what is going on? = observer comments

APPENDIX B

END OF EACH IMPLEMENTATION PHASE ITERATION INTERVIEW PROTOCOL

Thank you for taking the time to talk with me today. The purpose of this interview is to learn about your experiences using embedded formative assessment probes in science, specifically what has happened since the last probe was implemented. There are no right or wrong answers, or desirable or undesirable answers. I would like you to feel comfortable saying what you really think and how you really feel. If it's okay with you, I will be tape recording our conversation since it is hard for me to write down everything while simultaneously carrying an attentive conversation with you?

1. Could you please start off by sharing what probe you last embedded and the context of what students were learning about following the probe?
2. How did the data gathered from the formative assessment probe impact your teaching practices?
3. What were your goals for embedding this formative assessment probe?
4. What are your beliefs about your students' abilities for this current sequence of instruction?

I brought the Lesson Plan Binder you have been annotating. I would like to flip through and talk about some of these changes.

    a. Why did you make any decisions to change your planned instruction? (Evidence? Resource? Person?)
    b. How did you evaluate its success?
5. Were there any additional times that you used the probe supplemental materials outside of the collaborative sessions following the implementation.
6. Is there anything else that you would like to share with me about you experiences with embedding formative assessment probes?

APPENDIX C

END OF THE SCIENCE UNIT INTERVIEW PROTOCOL

Thank you for taking the time to talk with me today. The purpose of this interview is to learn about your experiences using embedded formative assessment probes in science, specifically your experiences over the course of this unit. There are no right or wrong answers, or desirable or undesirable answers. I would like you to feel comfortable saying what you really think and how you really feel. If it's okay with you, I will be tape recording our conversation since it is hard for me to write down everything while simultaneously carrying an attentive conversation with you?

Introductory Questions

1. Describe your current teaching situation. (consider probes – grade, subjects, types of students (level of ability), comfort – clarify what you mean or clarify that it is open to teacher perception)

The rest of the questions I have for you are only concerned with science instruction.

2. How do you maximize student learning in your classroom? (learning)
3. How do you describe your role as a teacher? (knowledge)
4. How do you know when your students understand? (learning)

5. In the school setting, how do you decide what to teach and what not to teach? (knowledge)
6. How do you decide when to move on to a new topic in your classroom? (knowledge)
7. How do your students learn science best? (learning)
8. How do you know when learning is occurring in your classroom? (learning)

9. Describe your experiences with embedding formative assessment? (Do you find it to be a valuable practice?)
10. How would you describe embedded formative assessment to someone who is unfamiliar with the concept? (to a teacher that is unfamiliar, to a policy maker, to a parent, to a student)
11. How would you define collaboration? (in general? Or in specific science contexts?)
12. Tell me about the nature of your collaborative relationship with your colleagues with whom you work on science units?
13. In your experience, what factors influence the nature of your collaborative relationship with your science colleagues? (Goals? Frequency? Productivity?)

The Semester Science Unit

The following questions all pertain to the ___________ science unit of study from the Fall.

1. Please provide a general overview of the unit.
2. What formative assessment probe(s) did you embed?
    a. Where did you embed formative assessment probes?
    b. Why did you choose the probe(s) you did?
3. Describe the (first/second/third) time you implemented the probe.
    a. How did you/the group make sense of the data collected? (Was it difficult? Surprising?)
    b. How comfortable were you personally in discussing the results?

c. How do you feel personal factors influenced your participation? (SMK? PCK? B&A?)
d. What were some of the take-aways for you from the collaborative session?

4. In the planning session(s) I noticed that you _________ can you tell me __________? (Inquire about patterns/noticings seen across the four implementation phases)

Closing Question

1. Considering the science unit and embedding formative assessment, what is a story that best represents your perception of successful formative assessment instructional outcomes? (does it have to be a personal narrative?)
2. Is there anything else that you would like to share with me about you experiences with embedding formative assessment probes?

APPENDIX D

START LIST OF CODES

Instructional Resources (IR)

- Curriculum (IR-C)
- Professional Development (IR-PD)
- Other Tools (IR-OT)

Teacher Resources (TR)

- Subject Matter Knowledge (TR-SMK)
- Pedagogical Content Knowledge (TR-PCK)
- Beliefs (TR-PCK)

Instructional Outcome (IO)

- Offload (IO-O)
- Adapt (IO-A)
- Improvise (IO-I)

BIBLIOGRAPHY

Allen, M. (2010). *Misconceptions in primary science*. Berkshire, England: Open University Press.

Amador, J. (2016). Mathematics pedagogical design capacity from planning through teaching, *18*, 70–86.

American Association for the Advancement of Science (AAAS). (2009). *Benchmarks for science literacy*. *Advancement Of Science*. New York: Oxford University Press. https://doi.org/10.1007/978-94-6209-497-0_11

Andrade, H. L., & Cizek, G. J. (Eds.). (2010). *Handbook of formative assessment*. New York: Routledge.

Arias, A. M., Bismack, A. S., Davis, E. A., & Palincsar, A. S. (2016). Interacting with a suite of educative features: Elementary science teachers' use of educative curriculum materials. *Journal of Research in Science Teaching*, *53*(3), 422–449. https://doi.org/10.1002/tea.21250

Atkin, J., Black, P., & Coffey, J. (2001). *Classroom assessment and the national science education standards*. Washington, US: National Academies Press.

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education*, *59*(5), 389–407. https://doi.org/10.1177/0022487108324554

Banilower, E. R., Smith, P. S., Weiss, I. R., Malzahn, K. A., Campbell, K. M., & Weis, A. M. (2013). *Report of the 2012 national survey of science and mathematics education*. *Horizon Research, Inc.*

Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice*, *18*(1), 5–25. https://doi.org/10.1080/0969594X.2010.513678

Beyer, C. J., & Davis, E. A. (2012). Developing preservice elementary teachers' pedagogical design capacity for reform-based curriculum design. *Curriculum Inquiry*, *42*(3), 386–413. https://doi.org/10.1111/j.1467-873X.2012.00599.x

Black, P., & Wiliam, D. (1998). Inside the black box : Raising standards through classroom assessment. *Phi Delta Kappan*. https://doi.org/10.1002/hrm

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, *21*(1), 5–31. https://doi.org/10.1007/s11092-008-9068-5

Brandon, P. R., Young, D. B., Shavelson, R. J., Jones, R., Ayala, C. C., Ruiz-Primo, M. A., … Furtak, E. M. (2008). Lessons learned from the process of curriculum developers' and assessment developers' collaboration on the development of embedded formative assessments. *Applied Measurement in Education*, *21*(931845722), 390–402. https://doi.org/10.1080/08957340802347886

Brown, M. W. (2002). *Teaching by design: Understanding the intersection between teacher practice and the design of curricular innovations*. Northwestern University.

Brown, M. W. (2009). The teacher-tool relationship: Theorizing the design and use of curriculum materials. In J. T. Remillard, B. A. Herbel-Eismann, & G. M. Lloyd (Eds.), *Mathematics teachers at work: Connecting curriculum materials and classroom instruction* (pp. 17–36). New York: Routledge. https://doi.org/10.4324/9780203884645

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

Driver, R., Squires, A., Rushworth, P., & Wood-Robbinson, V. (1994). Research into childrens ideas. *Making Sense of Secondary Science*.

Falk, A. (2012). Teachers learning from professional development in elementary science: Reciprocal relations between formative assessment and pedagogical content knowledge. *Science Education*, *96*(2), 265–290. https://doi.org/10.1002/sce.20473

Furtak, E. M., Kiemer, K., Circi, R. K., Swanson, R., de León, V., Morrison, D., & Heredia, S. C. (2016). Teachers formative assessment abilities and their relationship to student learning: findings from a four-year intervention study. *Instructional Science*, *44*(3), 267–291. https://doi.org/10.1007/s11251-016-9371-3

Furtak, E. M., & Ruiz-Primo, M. A. (2008). Making students' thinking explicit in writing and discussion: An analysis of formative assessment prompts. *Science Education*, *92*(5), 798–824. https://doi.org/10.1002/sce.20270

Furtak, E. M., Ruiz-primo, M. A., Shemwell, J. T., Ayala, C., Brandon, P. R., Shavelson, R. J., … Marie, E. (2008). On the fidelity of implementing embedded formative assessments and its relation to student learning, *7347*. https://doi.org/10.1080/08957340802347852

Gotwals, A. W., & Birmingham, D. (2016). Eliciting, identifying, interpreting, and responding to students ideas: Teacher candidates growth in formative assessment practices. *Research in*

Science Education, *46*(3), 365–388. https://doi.org/10.1007/s11165-015-9461-2

Heritage, M. (2007). Formative assessment: What do teachers need to know and do? *Phi Delta Kappan*, *89*(2), 140–145. https://doi.org/10.2307/20442432

Heritage, M., Kim, J., Vendlinski, T., & Herman, J. (2009). From evidence to action: A seamless process in formative assessment? *Educational Measurement: Issues and Practice*, *28*(3), 24–31. https://doi.org/10.1111/j.1745-3992.2009.00151.x

Hondrich, A. L., Hertel, S., Adl-Amini, K., & Klieme, E. (2015). Implementing curriculum-embedded formative assessment in primary school science classrooms. *Assessment in Education: Principles, Policy & Practice*, (November), 1–24. https://doi.org/10.1080/0969594X.2015.1049113

Keeley, P. (2011). *Uncovering student ideas in life science*. Arlington, VA: NSTA Press.

Keeley, P. (2014). *What are they thinking?:Promoting elementary learning through formative assessment*. Arlington, VA: NSTA Press.

Keeley, P., Eberle, F., & Farrin, L. (2005). Uncovering student ideas in science. Arlington, Va: NSTA Press.

Keeley, P., & Sneider, C. I. (2012). *Uncovering student ideas in astronomy: 45 formative assessment probes*. Arlington, VA: National Science Teachers Association.

Knight-Bardsley, A., & McNeill, K. L. (2016). Teachers' pedagogical design capacity for scientific argumentation. *Science Education*, *100*(4), 645–672. https://doi.org/10.1002/sce.21222

Luft, J. A., & Roehrig, G. H. (2007). Capturing science teachers' epistemological beliefs : The development of the teacher beliefs interview. *Electronic Journal of Science Education*, *11*(2), 38–63. Retrieved from http://ejse.southwestern.edu

Magnusson, S., Krajcik, J. S., & Borko, H. (1999). Nature, sources, and development of pedagogical content knowledge for science teaching. *Examining Pedagogical Content Knowledge*, 95–132. https://doi.org/10.1007/0-306-47217-1_4

McNeill, K. L., Katsh-Singer, R., & Pelletier, P. (2015). Assessing science practices: Moving your class along a continuum. *Science Scope*, *39*(4), 21–28.

Merriam, S. B., & Tisdell, E. J. (2016). *Qualitative research: a guide to design and implementation* (Fourth). San Francisco, CA: Jossey-Bass, a Wiley brand.

Miles, M. B., Huberman, A. M., & Saldana, J. (2014). Qualitative data analysis: A methods sourcebook. *Sage Publications, Inc*, (2014), 323–337. https://doi.org/10.1002/9780470699270

Mundry, S., Keeley, P., & Landel, C. (2009). *A leader's guide to science curriculum topic study*.

Washington: Corwin Press.

NGSS Lead States. (2013). Appendix F - science and engineering practices in the NGSS, (April), 1–33.

Ruiz-Primo, M. A., Furtak, E., Ayala, C. C., Yin, Y., & Shavelson, R. J. (2010). Formative assessment, motivation, and science learning. In H. L. Andrade & G. J. Cizek (Eds.), *Handbook of Formative Assessment*. New York: Routledge.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, *18*(2), 119–144. https://doi.org/10.1007/BF00117714

Sadler, P., Schneps, M., & Woll, S. (1989). A private universe. *Pyramid Film and Video, Santa Monica, CA.*

Saldana, J. (2016). The coding manual for qualitative researchers. New York: Sage.

Shavelson, R. J., Young, D. B., Ayala, C. C., Brandon, P. R., Furtak, E. M., Ruiz-Primo, M. A., … Yin, Y. (2008). On the impact of curriculum-embedded formative assessment on learning: A collaboration between curriculum and assessment developers. *Applied Measurement in Education*, *21*(4), 295–314. https://doi.org/10.1080/08957340802347647

Shepard, L. A. (2006). Classroom assessment. In R. L. Brennan (Ed.), *Educational Measurement* (pp. 623–646). Westport, CT: Praeger.

Sneider, C. I., Bar, V., & Kavanagh, C. (2011). Learning about seasons: A guide for teachers and curriculum developers. *Astronomy Education Review*, *10*(1), 19–40. https://doi.org/10.3847/AER2010035

Stake, R. E. (1995). *The art of case study research*. https://doi.org/10.1108/eb024859

Stavy, R., & Tirosh, D. (2000). *How students (mis-)understand science and mathematics: Intuitive rules*. New York: Teachers College Press.

Wiliam, D. (2010). An integrative summary of the research literature and implications for a new theory of formative assessment. In H. Andrade & G. Cizek (Eds.), *Handbook of Formative Assessment*. New York: Routledge.

Wiliam, D., Lee, C., Harrison, C., & Black, P. (2004). Teacher's developing assessment for learning: impact on student achievement. *Assessment in Education*, *11*(1), 49–65. https://doi.org/10.1080/0969594042000208994

Yen, C., Yao, T., & Chiu, Y. (2004). Alternative conceptions in animal classification focusing on amphibians and reptiles: A cross-age study. *International Journal of Science and Mathematics Education*, *2*(2), 159–174.

Yin, R. (2014). *Case study research: Design and methods*. *Essential guide to qualitative methods in organizational research* (Vol. 5). https://doi.org/10.1097/FCH.0b013e31822dda9e

Yin, Y., Shavelson, R. J., Ayala, C. C., Ruiz-Primo, M. A., Brandon, P. R., Furtak, E. M., … Young, D. B. (2008). On the impact of aormative assessment on student motivation, achievement, and conceptual change. *Applied Measurement in Education*, *21*(4), 335–359. https://doi.org/10.1080/08957340802347845

Yin, Y., Tomita, M. K., & Shavelson, R. J. (2013). Using formal embedded formative assessments aligned with a short-term learning progression to promote conceptual change and achievement in science. *International Journal of Science Education*, *0693*(January 2015), 1–22. https://doi.org/10.1080/09500693.2013.787556