Fall 2013

Elementary teachers' assessment actions and elementary science education: formative assessment enactment in elementary science

David Riley Pierson
University of Iowa

ELEMENTARY TEACHERS' ASSESSMENT ACTIONS AND ELEMENTARY SCIENCE EDUCATION: FORMATIVE ASSESSMENT ENACTMENT IN ELEMENTARY SCIENCE

by

David Riley Pierson

A thesis submitted in partial fulfillment of the requirements for the Master of Science degree in Science Education in the Graduate College of The University of Iowa

December 2013

Thesis Supervisor: Assistant Professor Cory Forbes

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

MASTER'S THESIS

This is to certify that the Master's thesis of

David Riley Pierson

has been approved by the Examining Committee for the thesis requirement for the Master of Science degree in Science Education at the December 2013 graduation.

Thesis Committee: ___________________________
Cory Forbes, Thesis Supervisor

____
Soonhye Park

____
Gabriele Ludewig

To my wife, Alisa, and Riley, Nicholas and Cara

“One of the hardest things teachers have to learn is that the sincerity of their intentions does not guarantee the purity of their practice.”

Steven Brookfield, *Becoming a Critical Reflective Teacher*

ACKNOWLEDGEMENTS

A special thanks to my thesis supervisor and academic advisor, Dr. Cory Forbes, for the guidance and support in researching and writing this thesis.

TABLE OF CONTENTS

LIST OF TABLES

TABLE

CHAPTER I
THE SIGNIFICANCE OF STUDYING FORMATIVE ASSESSMENT IN ELEMENTARY SCIENCE

Problem Statement

Currently described and enacted curriculum in K-8 classrooms is poorly designed for the purpose of building content knowledge according to analyses of science curricula in the United States (National Research Council [NRC], 2007). Decades of promoting inquiry-based science teaching and learning as a foundation of science education reform has failed as many elementary students are not given opportunities to engage in essential features of classroom inquiry (NRC, 2000).

Further, in light of the significant data on the effectiveness of formative assessment strategies on student learning (Heritage, 2007; Hermain and Choi, 2008; Furtak, 2008; Ruiz-Primo & Furtak, 2006; Ayala et al., 2008; and Furtak, 2006), elementary science educators' use of these strategies is not widespread. The differences in the use of these strategies by the elementary teachers that *are* enacting them have not been addressed by educational research as well. In particular, the actual usage of formative assessment strategies by elementary teachers has been paid little attention.

Purpose and Research Question

There is current research looking specifically at elementary science teachers and formative assessment implementation in their classrooms. *Reflective Assessment for Elementary Science in Iowa* (RAES-Iowa) is a three-year project designed around formative assessment use in elementary classrooms. Elementary science teachers are provided support through professional development over a three-year period in order to support formative assessment practices in their classrooms. More specifically, a 4-step formative assessment practice known as Reflective Assessment (RA) is being used by 3rd-6th grade teachers. Research has shown that elementary teachers' use of RA strategies has strong links to learning gains in science. There are four clear purposes of this project: a) to promote teachers' effective use of RA in their instruction and b) to increase teachers' content knowledge, c) to better engage elementary students in scientific practices and d) to promote student learning of scientific concepts. It is clear to see that this is a large project that spans many teachers, students and grade-levels.

In order to gain understanding of formative assessment use in elementary education this thesis study, which is situated within the RAES-Iowa project, focuses on the use of formative assessment in elementary classrooms where teachers are learning to implement Reflective Assessment. In a comparative case study of two elementary teachers enacting the same science curriculum, an Earth Science module on water, this thesis study seeks to address the following question:

1. How do elementary teachers think about and engage in formative assessment differently to support students' learning of Earth science concepts?

Significance

The significance of this question is to better understand how elementary science teachers enact formative assessment. A better understanding of how teachers think and engage in actions related to formative assessment can provide insight for further study on the implementation of formative assessment strategies by elementary teachers. It is my contention that formative assessment holds the key in transitioning teachers employing traditional methods of instruction towards research-based effective instructional strategies. Identifying the differences that exist among elementary teachers' resulting classroom actions related to their thinking of formative assessment is a critical first step for improving the use of this type of assessment in elementary science.

CHAPTER II
LITERATURE REVIEW

Introduction

Science education practices that take rise from strong research-based evidence on student learning are the gold standard for science education. In recent years formative assessment has emerged as a critical component of effective classroom instruction for science educators (Furtak, 2008; Heritage, 2007). Understanding the ways in which the use of formative assessment ultimately leads to increased student learning of science concepts is an important preliminary step in the adoption and implementation of formative assessment strategies by science teachers. In the following review of research several different critical aspects surrounding formative assessment are examined.

Beginning with research on how children learn is an important first step because ultimately it is this understanding that should guide strategies employed by teachers in developing increased student understanding of science concepts. This includes an understanding of constructivist theory and how inquiry science and other curricular developments, such as the BSCS 5E Model, were spawned out of constructivism research. The role of *engagement*, specifically, in science instruction will then be examined in relationship to student learning as well. These areas (constructivism, inquiry science, and the BSCS 5E model) all point to formative assessment practices as a critical feature when addressing student learning, so it is important to examine these individually to shed light on their connection to formative assessment.

Next, after establishing the rationales of formative assessment use through the constructivist lens, the role of formative assessment in meeting science standards set forth by this research is reviewed. The science standards, only established in relatively recent history, are reflective of research on student learning of science and therefore rely on formative assessment enactment as well.

After having established the research on student learning and the science standard development, a closer look at what formative assessment specifically entails is laid out. Defining formative assessment is important in order to move forward with the review on the benefits (i.e. impact on learning and motivation) and types of formative assessment. This then allows for a discussion of Reflective Assessment, which is the current, formative assessment-based teaching strategy the teachers in this comparative case study are using.

Lastly, I will lay out the connection between teaching and the use of formative assessment. This includes the research that supports the use of formative assessment in student learning increases. This gets at the heart of the purpose for this study in gaining a greater understanding of the importance of using formative assessment in science instruction to lead to increased student learning of science concepts.

Research on Learning

To get to the core of why formative assessment strategies are effective and well-grounded teaching strategies it is critical to review the research on learning that formative assessment is built on. To be clear, it is highly important to establish that formative assessment aligns with current research on learning. An understanding on how children learn is critical in designing and implementing instructional practices that are effective. Formative assessment is the focus of this review because it is based on current perspectives on how children learn, so it is essential to look at this particular research.

The idea that learning is a process that involves constructing internal mental representations and that these mental representations change in light of new understandings reflects a cognitive model of learning that is in contrast to the behavioral views of learning that have traditionally dominated school science.

Research on learners' previous experiences and background knowledge and its connection to learning new knowledge gets at the heart of an epistemology known as constructivism (Ultanir, 2012; Sandoval & Reiser, 2004; Cobb, Yackel & Wood, 1992). It was concluded in a review of constructivism's roots involving Piaget, Dewey and Montessori (Ultanir, 2012) that traditional forms of instruction involving memorizing hold a place in education, but do not create understanding. The traditionally behavioral approach for science instruction must dramatically shift toward the more cognitively oriented constructivist approach (Tennyson & Rasch, 1988).

Constructivism

Ultanir (2012), in a thorough examination of constructivism, interpreted three of the most well-known educational theorists' ideas. John Dewey, Jean Piaget and Maria Montessouri each set forth their views on how knowledge is constructed. The examination of these three constructivist theorists' ideas by Ultanir (2012) concluded, among many things, that one of the common threads amongst these theorists is that learners are not passive receptors of knowledge given by the instructor.

It is the building of knowledge that is emphasized in the constructivist approach, meaning that there is a starting point, so to speak, and that knowledge is built from that point for a person based on their experiences with their environment or social context, as opposed to passive reception (Ultanir, 2012).

Terhart (2003) describes the role of the teacher through the lens of constructivism:

> On the other hand, it is possible and responsible to understand teaching, and the practice of teaching, as something that makes stimulating environments available, which make things easier.

He goes on to explain that constructing these environments provides insights for students in developing their understanding and creates a situation in which independent learning can be facilitated. This is in direct contrast to his description of instructivism, which he explains as learners receiving information transmitted by teachers. This is identified by Terhart (2003) as "harmful and meaningless" in his study of constructivist didactics in Germany.

In a seminal publication on learning (The National Academies Press, 2000), Chapter 4 is devoted to research and understandings on how children learn. Jean Piaget, as mentioned earlier, is recognized as key theorist that emphasized the environmental stimulations that children receive that activate cognitive development. Piaget described cognitive stages that children go through, each involving significantly different cognitive schemes. This was significant due to the fact that it was in sharp contrast to the prior belief that a child's mind was like a blank slate, or *tabula rasa*. Piaget, as such, is viewed as pioneer of cognitive development of children and his work is highly influential in our current understanding of how children learn mainly because it established children as active learners.

In their conclusion of research on children's ability to learn, the National Academies Press (2000) reiterate that young children construct knowledge through various modes and their experiences are heavily involved in learning. Research on learning indicates that children are highly capable of reasoning, though their experiences and knowledge may be lacking. This means that their experiences only serve to enhance their learning as they refine and improve, as opposed to their minds being a blank slate that passively receives information. It is clear to see that identification of students' understandings is important for the teacher in creating learning experiences designed to build further understanding.

The National Academies Press (2000) publication "How People Learn: Brain, Mind, Experience, and School" point to the collaborative studies on learning environments that cognitive and developmental psychologists (along with educators) are involved in and the importance of these studies on the nature of learning and teaching. In essence, research continues to indicate that teaching practices that are aligned with the constructivist theories on learning are necessary if we are to improve science learning.

David Perkins (1999) describes constructivism as more than just one thing. He describes constructivism as multidimensional; *active learners, social learners,* and *creative learners*, in order to inform teachers how to organize learning experiences for their students. This only adds to further clarify the complex nature of constructivism and in understanding how different dimensions are involved in how children learn.

Along these same lines of active learning, the social environment was of particular interest of Vygotsky (1978) in his research on young learners. His *zone of proximal development* is iconic in the field of developmental psychology and provides a clear foundation for constructivist theory, and is subsequently the foundation of formative assessment. The zone of proximal development is described as a "bandwidth of competence." This essentially means that there is level that each learner is at (actual level) and that, provided the appropriate supports, there is a potential development level that can be reached.

Inquiry Science

One such approach to creating learning experiences is identified as inquiry science. This constructivist approach to science education is especially apt to encourage a variety of learning styles. The prominence of inquiry in science education reform is backed by numerous studies. Scientific inquiry includes a spectrum of abilities that scientists use to make understanding of the world around them (Yeh, Jen, & Hsu, 2012). A clear, positive trend favoring inquiry-based instructional practices has been indicated in a study performed by Minner et al. (2009).

To study the effectiveness of inquiry-based instruction (Wilson, Taylor, Kowalski, & Carlson, 2009) a laboratory-based, randomized control study was conducted and showed that students in inquiry-based classrooms reached significantly higher levels of achievement compared to the their peers receiving traditional instruction. This study contributed to a growing number of studies that provide evidence for inquiry-based instruction.

Undoubtedly there are many different perspectives on what constitutes "inquiry-based instruction." The NRC describes ''essential features of classroom inquiry'' (NRC, 2000, p. 25) including:

> (1) Learners are engaged by scientifically oriented questions.
> (2) Learners give priority to evidence, which allows them to develop and evaluate explanations that address scientifically oriented questions.
> (3) Learners formulate explanations from evidence to address scientifically oriented questions.
> (4) Learners evaluate their explanations in light of alternative explanations, particularly those reflecting scientific understanding.
> (5) Learners communicate and justify their proposed explanations.
> Inquiry is a central theme in the National Science Education Standards (NSES; NRC, 1996).

According to Piaget (1978) there will be a heterogeneous mix of what he described as concrete and formal operational staged students in any given classroom. This heterogeneity implies the implementation of hands-on student experience to better benefit all students. An exemplary classroom, therefore, would be a classroom that uses a lot of inquiry activities to allow students to construct concepts by various involvements with the concept. Accordingly, different kinds of inquiry science have been established upon the grounding in constructivist theory.

BSCS 5E Model

While the Biological Sciences Curriculum Study (BSCS) 5Es and inquiry are not equivalents, the former represents an instructional model based on constructivist theories of learning that gives a well-organized approach to teaching that promotes student inquiry (Wilson et al., 2010). The report put forth by Bybee et al., (2006) summarized recent research on the sequencing of science instruction, including laboratory experiences, to facilitate student learning. The report provides a rationale and empirical support for the BSCS 5E Instructional Model.

In a study of the effect of lessons applied in the subject of Inclined Projectile Motion, using the 5E Model as a basis, researchers reported success in achievement and motivation towards the subject when compared to traditional science instruction (Ergin et al., 2008). The National Research committee (NRC, 2006) proposed the phrase "integrated instructional units" in which "diagnostic, formative assessments are embedded into the instructional sequence and can be used to gauge students' developing understanding and to promote their self-reflection on their thinking." This speaks to the importance of sequencing science instruction for students and the critical nature of formative assessments in doing so.

These instructional units correlate directly to the 5E model (Bybee et al., 2006). The few studies that exist suggest that the BSCS Instructional Model is effective, or in some cases, comparatively more effective than alternative teaching methods in helping students master science subject matter (Akar, 2005; Coulson, 2002).

The 5 E effectiveness model: Origins and effectiveness (Bybee et al., 2006), have established an *engagement* stage as being a fundamental beginning for students in constructing science concepts. According to this model, *engagement* takes place when the teacher creates student interest, elicits students' questions, and deciphers students' prior knowledge with respect to the concepts to be taught (Goldston et al., 2010).

Based on the features of inquiry established described by the NRC, the importance of engagement of the learner is clearly of great value. Teaching strategies that engage students in learning through scientific investigations are more apt to increase conceptual understanding than are strategies that rely on more traditional techniques, which are often necessary in the current standardized-assessment laden educational environment (Minner et al., 2010). Thus, the importance of *engaging* students in processes of science (National Research Council, 2003) has emerged to the forefront of scientific literacy.

Engagement

Engagement, as it relates to the BSCS 5E Instructional Model, can be summarized by the teacher accessing the learners' prior knowledge, i.e. using formative assessment, and helping them become engaged in a new concept through the use of short activities that promote curiosity and elicit prior knowledge (Bybee et al., 2006). Engagement is the first step of the 5 E instructional model, in which studies (Boddy et al., 2003; Balci et al., 2005; Liu et al., 2009; Roehrig & Garrow, 2007) have shown increase in students' motivation to learn and achievement. The key is for students to mentally focus on an object, problem, situation, or event (Bybee et al., 2006).

The Iowa Core and the NRC (2007; 2011) promote scientific practices and students' engagement of science-as-inquiry as essential for students to learn science. It has been demonstrated in recent studies (Geier et al., 2008) that standards-based, inquiry science curriculum can lead to standardized achievement test gains when the curriculum is highly specified, developed, and aligned with professional development and administrative support. In a review of the studies of the impact of inquiry science instruction on K-12 student outcomes through the past twenty years Minner and colleagues (2010) found that clear positive trends favoring inquiry-based instructional practices existed. As also indicated by research, though, elementary science rarely focuses on conceptual understanding of science but instead emphasizes student engagement and hands-on activities (e.g. Beyer & Davis, 2008; Eshach, 2003; Forbes, Biggers, & Zangori, 2013).

Science Standards and Current Research

Science education reform across the nation now also has at its forefront science, technology, engineering and mathematics (STEM) education to improve the preparedness of students for life in the 21st century (National Research Council [NRC], 2007; 2011). The significance of STEM proficiency lies in data from the U.S. NAEP assessment that indicates that 34% of U.S. elementary students classified as "proficient" in science (Kuenzi, 2008). This is also concerning as U.S. students are being outperformed by students from other countries on international standardized assessments.

Recent new understandings of science, teaching and learning lead to a new nationally scaled effort of establishing a new framework for science education (National Research Council, 2012). The Next Generation Science Standards (Achieve, 2013) were recently released as a new set of science education standards. Drawing from a framework described as:

> …a broad description of the content and sequencing of learning expected of all students by the completion of high school…

developed by in *A Framework for k-12 Science Education: Practices, Cross-cutting Concepts and Core Ideas* (National Research Council, 2012).

The NGSS (Achieve, 2013) describe the seven conceptual shifts that make the NGSS new and different from the previous science standards. These conceptual shifts include the interconnectedness of the natures of science with the real world; performance expectations instead of curriculum; k-12 coherence; deeper understanding and application of content; integration of science and engineering; college, career and citizenship preparedness; and alignment with common core state standards.

The vision set forth in this framework is a very similar to the NSES scientific literacy focus, but with an added emphasis on engineering education. The committee authoring this framework set two major goals;

1) educate all students in science and engineering
2) provide foundational knowledge for those who will become the scientists, engineers, technologists, and technicians of the future.

Though one could argue that this might be labeled as scientific literacy, it does at least spell out more clearly the framework from which the NGSS were born from.

In the NRC framework (National Research Council, 2012) a thorough account of what science education should be guided by is laid out. In their description of attaining their vision the NRC describes three ways in which they will attempt to push science education towards coherence. The first way describes approaching learning as a "developmental progression", which reflects a cognitive approach to learning.

Along these same lines, the third way described explains the importance of the interconnectedness of knowledge and practice in designing learning experiences. The recent unveiling of the final draft of the *Next Generation Science Standards*(Achieve, 2013) is represents a shift of focus from what students know about science to what students understand. This more cognitive learning emphasis on the surface reflects some of the underpinnings of the use of formative assessment in the science classroom, and has huge implications for teachers' use of formative assessment in science as a classroom strategy to meet current science standards.

In an attempt to explain why elementary science education in the U.S. is underperforming, then, the focus on elementary science educators seems to be the logical first step. Elementary students have demonstrated abilities to engage in scientific practices to develop deep conceptual understandings of science, despite assumptions about their developmental abilities (e.g. Cavagnetto, Hand, & Norton-Meier, 2010; Hapgood, Magnussen, & Palinscar, 2004; Lehr, & Schauble, 2005; McNeill, 2011; Metz, 2004; 2008). The multi-subject demands that formal education systems put on elementary teachers often translates to poor subject matter knowledge in science (Anderson & Mitchener, 1994). Teaching inquiry-oriented science is especially complex (Crawford, 2000; Metz, 1995) and typically requires teachers to teach in a way that is different from how they were taught (Windschitl, 2003).

Formative Assessment Defined

It is clear that researching classroom practices that correlate with increased science proficiency guided by science education and STEM standards are crucial for the future of elementary science education. One such classroom practice involves identifying current student understanding, identifying the gap between that understanding and established proficient understanding, and strategizing how to bridge that gap. This is a process generally known as formative assessment. Formative assessment strategies, such as Reflective Assessment, have been shown to be easily implemented in elementary classrooms and takes little additional time for teachers (Kennedy, Long, & Camins, 2009; Osmundson, Dai, & Herman, 2011).

Teaching strategies that engage students in learning through scientific investigations are more apt to increase conceptual understanding than are strategies that rely on more traditional techniques, which are often necessary in the current standardized-assessment laden educational environment (Minner et al., 2010). Thus, the importance of *engaging* students in processes of science (National Research Council, 2003) has emerged to the forefront of scientific literacy. Reports in recent years, such as *The 5 E effectiveness model: Origins and effectiveness* (Bybee et al., 2006), have established an *engagement* stage as being a fundamental beginning for students in constructing science concepts. According to this model, *engagement* takes place when the teacher creates student interest, elicits students' questions, and deciphers students' prior knowledge with respect to the concepts to be taught (Goldston et al., 2010). The *engage* phase aligns with the current research on brain-based learning that advocates mental engagement of students to elicit what the learner already knows and values and how this information connects to the new learning experiences (Goldston et al., 2010).

Strategies for engaging students would seem to be key component in an inquiry-based classroom. Black and Wiliam (1998a) explain an assessment strategy known as formative assessment as "all those activities undertaken by teachers, and/or by their students, which provide information to be used as feedback to modify the teaching and learning activities they are engaged." This category of assessment (formative) can be thought of as "assessment *for* learning and not *of* learning"(Black and Wiliam, 1998b; Pellegrino, Chudowsky, & Glaser, 2001). Formative assessment is a process through which assessment-induced evidence of student learning is collected and instruction is adapted in response to this feedback (Cauley and McMillan, 2010).

Benefits of Formative Assessment

It is important here to look at the benefits that are specific to the use of formative assessment. Up to this point I have only referred to the roots of formative assessment in constructivist forums and have defined specific forms of formative assessment that briefly refer to their successes. The impact of formative assessment on learning and motivation for students are the two characteristics that get at the heart of why this teaching strategy should be adopted and implemented by science educators.

Impact of Formative Assessment on Learning

Recent publications, such as Taking Science to School (National Research Council of the National Academies, 2007), have indicated that research on the effectiveness of formative assessment across many school subjects implies significant results. Black and Wiliam (1998a) reviewed over 250 different articles centered on aspects of formative assessment and found an effect size between .4 and .7 in learning gains based on pre and post measures of student learning. It is important to note, to put things within a larger context, few strong empirical studies on formative assessment specific to science existed at the time of Black and Wiliam's study. In one such study (White & Frederikson, 1998) an examination of peer and self-assessment and its relationship to the development of understanding of scientific inquiry was conducted on four middle school science classes. The results of this study indicated that students engaged in the reflective assessment process performed better on both project work and unit exams.

Impact of Formative Assessment on Motivation

Claims have also been made that formative assessment techniques bring benefits in terms of self-esteem and motivation, though evidence in this area is limited (Miller and

Lavin, 2007). Boddy et al. (2003) describe engagement of a student in lessons as depending upon their personal motivation. Karlsson (1996) describes motivation as "the development of conditions promoting intention to act or learn." The implication here is that if somehow formative assessment leads to increased motivation, and motivation is key in engaging students in learning, then formative assessment may ultimately be an instructional practice at the center of improved science learning.

Research shows a positive relationship between formative assessment, student motivation and achievement on both classroom and large-scale assessments (Cauley & McMillan, 2010). Unfortunately, in our current accountability environment, assessment is seen as a tool solely for measuring what students have learned for purposes of ranking students and schools (Heritage, 2007). In cases where the value of formative assessment is understood and is implemented through professional development, fidelity of implementation seems to have an impact on learning gains (Furtak et al., 2008). In such cases, greater fidelity of implementation leads to greater learning gains. In other examples where formative assessment was implemented wholesale across a district, success was reported on summative evaluations of learning (Pijanowski, 2008). It appears as though the focus should shift towards how to raise fidelity of implementation and why formative assessment isn't a larger part of science education practices in our schools.

Types of Formative Assessment

In reviewing studies on formative assessment; Heritage, 2007; Hermain and Choi, 2008; Furtak, 2008; Ruiz-Primo & Furtak, 2006; Ayala et al., 2008; and Furtak, 2006, establish formative assessment as being beneficial in assessment for learning in a science classroom. Formative assessment practices, according to *Taking Science to School* (National Research Council of the National Academies, 2007) frequently takes three distinct forms; on-the-fly, planned-for, and curriculum-embedded. *On-the-fly* formative assessment is an interactive formative assessment which focuses on gathering information about student learning whenever possible, in any student-teacher interaction (Ruiz-Primo & Furtak, 2006). *Planned-for* formative assessment is centered on some level of on-purpose planning on the part of the teacher, such as questioning strategies at a specific point in a lesson (Furtak, 2008). *Curriculum-embedded* formative assessments are assessments embedded at critical junctures in the on-going curriculum to intentionally create "teachable moments" (Furtak, 2008).

Shavelson & SEAL (2003) describe a continuum of formative assessment that involves informal formative assessment on one end to formal formative assessment on the other. In this continuum, planned-for formative assessment lies directly in the center between on-the-fly at left and curriculum-embedded at the right.

Planned-for Formative assessment

Planned-for formative assessment involves some level of deliberate planning on the part of the teacher (Furtak, 2008). Planned-for questioning strategies can be an effective tool for formative assessment (National Research Council of the National Academies, 2007). Teachers decide ahead of time how they will elicit students' thinking during the course of instruction during planned-for assessment (Heritage, 2007). When planned deliberately, assessment conversations, such as daily classroom talk, allow teachers to recognize students' conceptions during an activity already occurring in a classroom.

For example, in anticipation of a common misconception about a particular science concept during an investigation, a teacher has planned questions that elicit student discussion of this particular concept to gauge student knowledge. Certainly in an investigation about weight and mass, a teacher might plan a question relating to gravity to gain insight into how well students understand the relationship between weight, gravity and mass.

Curriculum-embedded Formative Assessment

Teachers or curriculum developers may embed assessments in the ongoing curriculum to intentionally create “teachable moments” (Furtak, 2008). According to Furtak (2008), these assessments are embedded at critical junctures, and designed so that feedback on performance to students is immediate and actions are taken by the teacher to close whatever deficiencies exist. In essence, feedback on problem areas revealed by the assessment is designed to address deficient areas. Diagnostic, formative assessments embedded into the instructional sequence can be used to gauge students’ developing understanding and promote their self-reflection on their own thinking as well (Bybee et al., 2006). Furtak et al., (2008) explains that embedded formative assessments are formal prompts inserted into curriculum that are designed to help teachers check understanding at key points during instruction and reflect on the next steps to needed to move students forward in their learning.

On-the-fly Formative Assessment

On-the-fly formative assessment is a more informal or interactive formative assessment which focuses on gathering information about student learning whenever possible, in any student-teacher interaction (Ruiz-Primo & Furtak, 2006). In studies where this type of formative assessment was under investigation (Ruiz-Primo & Furtak, 2006), teachers who employed more discussions, asked more concept-eliciting questions, and had the greatest diversity of strategies that used information gained about student understanding had the highest gains in academic performance.

In programs that have implemented informative formal assessment measures (Gallagher, 2009) trends have been identified and reported that indicate a shift from teaching for coverage to learning for understanding, among other positive effects on student learning. In such programs, the strategy of asking appropriate and well-thought-out questions on the spot, so to speak, have greatly enhanced teachers’ knowledge of what students think and understand (Crumrine & Demers, 2007).

In essence, on-the-fly assessment occurs spontaneously during the course of a lesson to identify what misconceptions are present for students that might need to be addressed (Heritage, 2007). Teachers may overhear a small group discussion about an intriguing question that clearly identifies a misconception the group has about the concept in question. At that moment, the teacher may address this misconception that they may not have otherwise been able to identify through traditional instruction until summative assessment (Heritage, 2007). This technique enables teachers to address misconceptions by changing the direction of the lesson and engaging the student to think about their conceptions.

According to Bybee et al. (2006), the activity should make connections between past and present learning experiences, expose prior conceptions, and organize students' thinking toward the learning outcomes of current activities. Clearly the benefits of this identification, and the adjustment of lesson planning as a result, would logically lead to a greater understanding of science concepts for students. Teachers can increase student motivation by facilitating enjoyable learning experiences and allowing students to make decisions about their learning. This increased student motivation increases their learning because the student is actively engaged in the learning experience (Boddy et al., 2003).

Reflective Assessment

Reflective assessment (RA) is a 4-step strategy rooted in formative assessment. Essentially there is as planning stage labeled "anticipate" where teachers design a way to gather students' misconceptions about a science concept, or more specifically to learn the students' current understanding. In the next stage, labeled "review", the teacher is identifying the misconception/current understanding. The third stage, or "reflect", the teacher assesses the current state of the class based on what was reviewed. The "adjust" stage is the final stage in which lessons are designed for instruction to provide constructive feedback to students about misconceptions.

In the *anticipate* step, the first step of reflective assessment, teachers are focused on one or two concepts in the given lesson. Teachers think about prior experiences with the concept with former classes and take these experiences into account when planning for the upcoming lesson. Identifying the concept and the student work that will be used to examine in order to determine the level of understanding of the class is the defining characteristics of this step.

In the next step of reflective assessment, known as *review*, teachers are getting a count of how many students "get it" and how many students "still need help." This is typically done with computer software for convenient data entry and retrieval or can be done by the teacher keeping a tally of the students who understand and those who do not. This review of student work by the teacher provides the teacher with the evidence they need to make the appropriate instructional decisions. The key characteristic of this step is a number comparison of students who understand the concept to students who still need help.

The instructional decisions made by the teacher, called the *reflect* step, involves the teacher identifying misconceptions that are held by students based on the work they submitted. The distinguishing aspect of this step is the teacher recognizing what the reviewed work shows about students' ideas of the concept of focus.

The last step of reflective assessment involves the next-step planning of the teacher. Referred to as the *adjust* step, this component of reflective assessment is characterized by what actions the teacher decides to take to address the misconceptions of students. The key factor to take into account by teachers in this step is the number of students who still need help. If relatively few students still need help then the teacher may choose actions that only apply to those students to address their misconceptions. Conversely, if a large number of students hold the same misconception about the concept then next-step actions of the teacher might involve a follow-up lesson, involving one of several strategies, to address the misconception.

Trauth-Nare and Buck (2011) used reflective assessment as a part of their action research on formative assessment. They found that reflective practices better informed formative assessment actions and teachers saw the value in reflective practices in their formative assessment actions despite low efficacy at the beginning of their study. This action research gives insight into the potential for currently practicing teachers in their use of reflective practices.

The effects of reflective assessment use by students on their inquiry practices in science were the focus of a study by Toth, Suthers & Lesgold (2002). They found that reflective assessment rubrics used by students greatly improved students' reasoning during inquiry activities. Though this form of reflective assessment is being used by students in their investigations, it provides support for much the same components of reflection used by teachers of their formative assessment data.

In the reflective assessment most relevant to this proposed thesis study, Kennedy, Long & Camins (2009) explain that when teachers focus their formative assessment use student learning can be increased. They refer to this as "guided teacher reflections" related to decision-making in classrooms. They have laid out a *Reflective Assessment Technique* which was developed through two National Science Foundation projects. This technique involves *anticipation*, *review*, *reflect*, and *adjust* as its steps for systematically implementing formative assessment actions. In their initial studies they have found clear positive trends for classrooms employing the reflective assessment technique.

Teachers and Formative Assessment

Some empirical evidence suggests that, in general, formative assessment leads to an increase in learning gains (Ayala et al., 2008), but how these formative assessments are designed, developed, embedded, and eventually implemented by teachers continues to be researched. The success of on-the-fly formative assessments relies more heavily on the teacher's "toolbox" of strategies that are implemented in the daily classroom discourse.

As compared with the curriculum-embedded and planned-for assessments, the on-the-fly formative assessment strategies are more heavily reliant on the actions of the teacher on a daily basis in the classroom in determining the direction of each unit/lesson based on the information he/she is receiving through on-the-fly formative assessment strategies. As the name implies, these actions are "on-the-fly", but nonetheless are important in determining the direction of the lesson and gather important pieces of information in lesson-planning.

What research hasn't shown is how decisions on next-steps are made by teachers once formative assessment data has been collected. Elementary teachers' decisions of what takes place in their classroom undoubtedly are driven by many factors. For instance, the impact of their content knowledge of the science content they are teaching is hypothesized to have a significant impact on what next-step measures are employed. To clarify, once a teacher has identified current levels of understanding of a concept by their students their own conceptual understanding of that concept might be a driving force in what the teacher decides to do next.

Another factor that might influence next-steps decision-making is the teachers' perspective on how students learn. If the teacher aligns with a more traditional view of student learning, as opposed to a constructivist view, then the next-steps might be reflective of this view.

The study of elementary teachers' described perspectives and enacted classroom decisions is especially interesting due to the differences in how elementary science teachers are prepared in contrast to higher levels of science education. The demands on elementary teachers to teach multiple subjects throughout their day, as opposed to secondary levels which likely teach science only, leads to the belief that elementary teachers' uses of research-based teaching strategies in science might be lacking. Though there is substantial research on the effectiveness of formative assessment for student learning, little attention has be paid to how elementary teachers use formative assessment in their class rooms. A closer look at the factors influencing implementation of formative assessment by elementary science teachers might expose important aspects for further research on the implementation of effective formative assessment. More importantly, by learning and implementing formative assessment strategies teachers previously implementing traditional instructional models can see more clearly the benefits for student understanding that research-based models of instruction can bring.

Conclusion

The research of constructivist approaches to teaching relies on creating opportunities, or environments in which learners can take in experiences in order to build understanding of a concept. The identification by the teacher of a student's current level of understanding is critical in planning and enacting these opportunities for students and can serve as a critical first step in becoming a constructivist teacher. In reviewing the research of the various components that surround formative assessment, two conclusions emerged prominently.

The first conclusion that can be made is that formative assessment has a clear, important role in science education as we move forward in teaching k-12 students the essential skills and concepts of science. In general, student work that provides evidence of their current level of understanding, formal or informal, should be used by educators in planning science instruction that is meaningful in building student understanding. Research indicates that students construct knowledge through specifically designed learning experiences, and that inquiry science and curricula implementation of models such as the BSCS 5E model increase student engagement needed for advancing their understandings in science. The essential skills and concepts identified are best developed through classroom practices rooted in this research in meeting the science standards that have been established. In fact, the newest standards specifically indicate constructivist practices in science education, thus making formative assessment essential.

The second conclusion that can be made is that stakeholders in science education, i.e. national policymakers, state and local administrators, parents, teachers, etc., must recognize the benefits of constructivist learning and take action towards supporting constructivist curricula training and enactment in k-12 science education. There is a clear desire for raising achievement in science of k-12 students, but the implementation of constructivist-rooted teaching strategies, such as formative assessment strategies, still appears to generally lacking.

Moving forward research must continue in how teachers adopt and enact formative assessment strategies. What does exemplary formative assessment enactment look like? What are teachers' perceptions of formative assessment and how do these perceptions affect their own use of formative assessment. Are more elementary science teachers using formative assessment strategies or constructivist-based strategies with their students now in light of this research? These are significant research questions moving forward in establishing constructivist enactments of science lessons. Changing the paradigm from the traditional idea that students are passive receivers of knowledge to constructivist thinking where student experiences guide their learning is critical in raising achievement in science. Teachers that proficiently formatively assess and design rich learning experiences accordingly are crucial for the future of science education.

CHAPTER III
RESEARCH METHODS

This study centers on the formative assessment practices of two elementary science teachers. The qualitative nature of studying these two teachers to gain a better understanding of how elementary science teachers engage in formative assessment practices differently has been sparsely examined prior to this study. This chapter explains the context of this study and describes the research design employed. The theoretical basis of qualitative research and case studies is then described as well as the data collection, coding and analysis of the data collected from these teachers. A description of the two case study teachers and their students is provided as well.

Study Context

In this section I describe the context of which this study is situated in order to give a better account of the nature of this study. A description of this study, which is embedded in a larger project, is given to provide clarity on how the two case study teachers were selected and a more specific account of these two teachers.

The RAES –Iowa Project

RAES-Iowa is a three-year study that involves 38 third through sixth grade teachers in learning to use Reflective Assessment, a four-step formative assessment strategy developed by the Lawrence Hall of Science. The project serves public school districts in the east-central region of a single Midwestern state. Two of the school districts involved are designated as "high needs" meaning that less than 85% of the 4^{th} grade students in these districts are deemed proficient in science (and greater than or equal to 20% of students qualify for Free/Reduced Priced Meals). The other school districts involved, including the district of the participants in this study, do not meet this designation. Full Option Science System (FOSS), Science Technology and Society (STC), and Insights science curriculum modules are used by the teachers in the RAES project at the third through sixth grade levels. These modules specifically involve three disciplinary content strands; Earth science, biological sciences, and physical sciences. The first year of the project the Earth science content strand was the focus, and this was the content strand used by the teachers in this study. During the 2012 – 2013 school year, the participating teachers in the RAES-Iowa study implemented one of the three science modules: Water, Earth Materials, or Landforms.

A professional development program to support teachers' learning to use the reflective assessment strategy (Professional Development for Reflective Assessment – PD4RA) was developed by science educators and curriculum developers at the Lawrence Hall of Science. The teachers of the RAES project and of this study participated in the PD4RA through facilitated collaborative teacher learning teams (CLTs) during the 2012-2013 school year. The CLTs met for an hour and a half bi-monthly after school for a total of 30 hours. These CLTs were on-going during this study. The focus of the CLTs was on teachers' implementation of the reflective assessment strategy in their Earth Science module. CLTs examined student artifacts, teacher logs, and video-recorded enacted lessons to reason collaboratively about students' ideas and misconceptions, unit activities, and next-steps strategies.

In addition to the after school meetings, two three and one-half hour workshops were held on Saturday mornings during the year for all RAES project participants. During these module-specific trainings teachers performed module investigations, anticipated and located likely points in modules for student misconceptions, and planned for use of specific next-steps strategies at particular points in the modules. The teachers also spent seven days during the summer of 2012 learning about Reflective Assessment and the Earth Science content modules.

Research Design

Qualitative Research

In order to answer this study's research questions a case study research approach is appropriate in many ways. If we define qualitative research as research that attempts to make sense of or interpret phenomena (Denzin & Lincoln, 2005) in the natural setting, then exploring these two teachers use of formative assessment in their class room qualitatively is certainly logical. The natural setting is a major aspect of qualitative research, so studying the two case study teachers in the natural setting of the problem at hand, differences in use of formative assessment, is most appropriate accordingly.

There are several characteristics of qualitative research that give it merit (Creswell, 2006). Natural setting as the source of data, researcher as key instrument of data collection, involving a holistic account, and a focus on participants' perspectives are just a few of these characteristics. Accordingly, these, and other characteristics of qualitative research are present in this study. This provides a justification for studying these teachers' differences in enactment of formative assessment qualitatively. Paying particular attention to the characteristic of "holistic account", this study attempts to identify and describe the many factors involved in a situation and tries to paint the big picture that emerges from the study (Creswell, 2006).

Case Studies

Qualitative case studies are a common nonexperimental alternative to experimental methods for answering questions about a causal hypothesis, especially in the academic fields of education and anthropology qualitative case studies (Shadish, Cook & Campbell, 2002). Lincoln & Guba (1985) go even as far as to state that case studies might even replace experiments of "causal-sounding" questions, though this is not widely accepted among qualitative theorists. Qualitative methods, such as case study, have the potential, though, to reduce enough doubt about causation to meet the needs of the stakeholders involved. Though causation is not the central focus of this study, it is clear that a comparative case study is a fairly widely accepted method in certain fields of research.

Oliver (2004) describes case studies as follows:

1. Does not control or manipulate variables
2. Studies phenomena in the natural context
3. Studies phenomena at one of a few sites
4. Uses qualitative tools and techniques for data collection and analysis

In looking at organizational culture, more specifically record-keeping in different cultures, Oliver (2004) used a comparative case study of Hong Kong, Australia, and Germany. In this study the focus of the case study was the phenomena (record-keeping) and the cultural context (Hong Kong, Australian, and German culture). Phenomena in diverse settings allowed for comparison of findings in order to gain a better understanding of different cultures. It is important to note, though, that generalizations cannot be made to whole populations from case study findings *only* (Shadish, Cook & Campbell, 2002).

Karen Rajaona Daka (2011) performed a case study in which the electrification of homes and children's education was studied in northern Madagascar. In this study survey data of 100 homes in a particular region of northern Madagascar was collected, among other data, to shed light on the connection between homes with (or without) electricity and the impact on the education of children living in those homes. Many descriptive statistics were used in this study to help researchers further understand the importance of having electricity in homes of school children.

In general, case study approaches in educational research have served to further inform empirical research. This study is thus well-grounded within this context to provide further qualitative data within the context of the RAES-Iowa project.

Data Collection

During the 2012-2013 school year the two teachers in this study implemented the FOSS Water unit. Data for this study was collected from three main sources; online teacher logs, interviews, and video-recorded lessons. The data collected from these three sources is necessary for gathering evidence to answer the two research questions of focus for this study. A description of the data collection is provided in the following sections.

Online Teacher Logs and Interviews

The *online teacher logs* were completed during the module enactments. The case study teachers completed three online teacher logs each. This log is organized into the 4 critical components of RA; *anticipate, review, reflect,* and *adjust* instruction. This is a self-reporting instrument that allows for consistent and ongoing measurement of teachers' perceived implementation of RA in their classroom. Kennedy, Long & Camins (2009) effectively used this instrument in prior research. Teachers completed multiple logs for each science unit they taught, including their Earth science units.

The two case study teachers of this study each completed three online teacher logs for the Earth science module on "water." The online teacher log has a total of six questions that that were scored by two partner scorers in the RAES project, which provided quantitative data further described later. The online teacher log consists of six questions that relate to the steps of RA. These six questions allow the teacher to express their use and perceptions of the steps or RA. For instance, question three of the online teacher log asks:

> What is the key concept you are assessing in the student work? What student work are you examining? What are you looking for in students' responses as evidence?

For this study, collecting this data and coding it according to the steps of RA provides us evidence of formative assessment perceptions that can then be qualitatively analyzed.

Qualitatively coding the four steps of Reflective Assessment based on the interview responses provides insight into the differences and similarities of the teachers' understandings and perceptions of these steps as well. The interviews were conducted using a protocol of formal questions regarding RA for a given Earth Science module lesson. The following are the questions that were asked:

a. What was the key concept you chose to look at for this log?
b. Where did that key concept come from? What makes it a good key concept to look for?
c. What type of student work did you look at for this RA cycle?
d. Why did you choose that particular student work sample?
e. What did you notice as you looked through the students' work?
    i. What did you see as evidence of student understanding?
    ii. What were examples of student misconceptions?
f. I can see from your log that you chose _______ as your next step strategy. Tell me how that went.

As is shown, each step of RA is addressed in the questions and provided the teacher an opportunity to not only explain what they did, but how they did it and why in some cases. This provided insight into the teachers' perceptions and rationales as well as verification of the use of each RA step.

Video

The video-recorded enacted module lessons were taken from each of the 2 case study teachers. Four full lessons were recorded for each teacher, which provided an opportunity to observe and compare the case study teachers' actual implementation of RA actions, and most of the videos were recordings that displayed the last stage of the four stages of RA. The last stage of RA is the "adjust" stage in which teachers design lessons to provide constructive feedback to students about misconceptions.

For clarity, these video-recorded enacted module lessons mostly provide visual evidence of this fourth step of RA, which, again, provided an opportunity to gather qualitative comparisons of the 2 teachers and help triangulate observations made from the online teacher logs and teacher interviews.

The first lesson was recorded in early November of the school year, and the fourth video was of a lesson given in early December. Though each of these videos provide for the qualitative analysis of the teachers' enactment of the steps of RA, they are primarily recordings of lessons that implement next-step actions, or the *adjust* step of RA. The third video sheds light on what the differences in implementing next-steps actions looks like for both of the case study teachers. This is mainly because each teacher gave the same lesson on the same day covering the exact same concept relating to water, but the lessons appear much different.

Case Study Data and Analysis

As mentioned, this is a case study of two elementary science teachers. The data collected is mostly qualitative data surrounding the various aspects of the implementation of formative assessment. This exploratory approach to studying teachers' differences in implementing the steps of Reflective Assessment for the same Earth Science content module and what these differences look like are the key focuses of this study. To do this, two teachers were selected that are implementing the same Earth Science content module focused on various concepts of water at the 3rd-grade level. In fact, much of the time they taught the same Earth science module lessons on the same days. The two teachers were also selected because both teach at the same school and are veteran teachers. The teachers, described as teacher A and teacher B in the study, have 23 years and 25 years respectively of teaching experience. Interestingly, though, their online teacher log scores, teacher knowledge scores, and student scores were much different.

The qualitative data comparisons of these two case study teachers could potentially shed light on the impact that varying levels of enactment of formative assessment actions have in elementary classrooms. The two research questions this study seeks to answer are significant because formative assessment use by elementary teachers has been paid little attention by educational research, especially specifically looking at the differences in how formative assessment actions are implemented. This could inform further research concerning how elementary teachers engage in effective use of formative assessment in science education differently and could potentially raise questions of causation that could warrant further study.

Participants and Context

Science Content Knowledge

Teachers' science content knowledge was the 2nd objective of the RAES-Iowa project. All of the 38 teachers were given an assessment of science content. The Center for Research in Mathematics and Science Teacher Development (CRMSTD) tested and validated three versions of the Diagnostic Science Assessment for Middle School Teachers, one for each of Earth, biological, and physical sciences. This assessment was given at the beginning of the summer of 2012 and again at the beginning of the 2013 summer institute. The students also are evaluated in this study according to their conceptual understandings of science. Student scores on pre-/post-module assessments were collected during the 2012-2013 year. In describing teacher A and teacher B in the following sections, the scores of their students on pre-/post-module assessments are given among other characteristics.

Teacher A teaches 3rd grade at an elementary school in a Midwestern state. There are 18 students in Teacher A's classroom, eleven boys and seven girls. On the pre-test given for the water module, Teacher A's students had an average score of 8.35. This test consisted of 37 points and was scored by a University of Iowa Science Education graduate student. This same student scored all of the student work in order to establish consistency in scoring. This same student also scored the posttests. The students in

Teacher A's class scored an average of 15.72 on the posttest, which was a gain of 7.37 from pre-test to posttest.

There is also quantitative data to consider for teacher A. The teacher content knowledge exam questions given to the participants in the RAES-Iowa study were categorized according to the content area addressed in the question (life, Earth, or physical science). Teacher A's score on the Earth science questions of this content knowledge exam were analyzed because the module being taught in this study was the FOSS water module, and Earth science topic. Teacher A answered 8 of the 12 Earth science questions correctly.

The online teacher logs collected from the 2 case study teachers were scored by two University of Iowa Science Education Department graduate students using a method known as partner scoring. For each log each partner scored the log using an "RA Log Rubric" that scores each of the 6 questions on the log on a 0 to 4 scale, with a 4 indicating the highest valued score for each questions. A 4 means that the teacher fully met the key concept addressed in the question in the answer they submitted in their log, whereas a 0 indicates "no response."

The score for each question on the online teacher log was coded and then an average core was established for that teacher by that partner scorer. Partner 2 then scored the same log for the same teacher and, too, established an average for that teacher on that same online teacher log. After each of the 2 partners established an average score for a teacher's log, the 2 averages of the partners were averaged to give the teacher an overall log average on that particular log. For the 2 case study teachers, 3 logs were submitted and coded in this manner. An average, then, was then calculated for their 3 log averages to give each teacher an overall teacher average based on their 3 logs.

Teacher B also teaches 3rd grade at the same elementary school as teacher A. In fact, teacher A and teacher B often collaborate in lesson planning and deliver lessons over the same content on the same days. Teacher B has 20 students in her class, 10 boys and 10 girls. On the pre-test given over water, scored as previously mentioned, teacher B's students scored an average of 7.33. On the posttest given for this unit, teacher B's students scored an average of 14. This was a pre-posttest gain of 6.67.

On the teacher content knowledge exam, teacher B's answer were analyzed in the same manner as teacher A. Teacher B answered correctly 7 out of the 12 on the Earth science questions. This score compared to the 8 out of 12 on the Earth science questions answered correctly by teacher A generally establishes at least very similar content knowledge in Earth science as measured by this content knowledge exam.

It can be concluded that teacher A and teacher B exhibit many similarities in the areas of content knowledge of Earth science, school environment, class size, gender count in their classrooms, and even years of experience teaching. There does exist, though, a .7 of a difference of log score average their online teacher logs of RA critical components for their module lessons over the same science content. The interesting piece of data to consider is that teacher A has the slightly lower RA online teacher log score, but had slightly higher student pre-posttest gains.

Table 1: Comparison of Case Study Teachers

Teacher	A	B
Years experience	23	25
Grade level	Third	Third
Number of Students(B,G)	18 (11, 7)	20 (10,10)
Students Pre-test score	8.35/37	7.33/37
Students Post-test score	15.72/37	14/37
ES Question Score	8	7

Data Coding

For each of the data sources a coding system was used to identify the four steps of reflective assessment. Any time one of the four steps of reflective assessment was described or enacted by the case study teacher a corresponding code was tagged accordingly. This coded data was collected throughout the analysis of each of the data sources. In this section I will describe how this coding system was created and applied for each of the data sources. Figure 1 (below) displays the codes and descriptions for each code that was used when collecting data.

Table 2: Reflective Assessment Codes and Descriptions

Code	**Description**
An	*Anticipate.* Concept and student work identified by teacher and any expectations by the teacher based on their experience
Re	*Review.* Teacher identifies a number of students who “get it” or “don’t get it” based on review of student evidence collected from student work
Rf	*Reflect.* Teacher describes misconceptions or what was learned from looking at evidence collected from student work
Ad	*Adjust.* Teacher describes/enacts next-steps actions to address student misconceptions

When the online teacher log data was collected, this data chart was created to organize this data according the four steps of Reflective Assessment. This data chart was organized in such a way that the step RA, (*An, Re, Rf*, *Ad*) was listed on the left of the data chart, and the description of the teacher action/expression of this step was written to the left of this code. This coding was used as the data from each online teacher log from each of the case study teachers was collected. There were three online teacher logs submitted from both teacher A and teacher B, therefore there were three coded data charts for each teacher that allowed for comparison between the two teachers of their actions/expressions of the four steps of RA.

Once data was collected and coded accordingly, the coded data was organized onto a spreadsheet to allow for comparison across the data points (logs, interviews, and videos). For example, teacher A's online teacher log data was collected and coded onto a data chart. The data from each log for teacher A was coded and recorded onto a separate data chart, thus three different data charts for teacher A were created. These three data charts each organized teacher A's data based on the four steps of RA. Next to each RA step (*An, Re, Rf, Ad*) code a description of the data that reflects that specific code is given. To reiterate, for each coded step of RA on all of the data charts a description of the data that reflects that code is also given.

Teacher B's online teacher log data was collected and coded in the same manner as teacher A's. This allowed for careful comparison of the online teacher log data between the two teachers, more specifically the comparison of the descriptions for each code. For example, the description for the *Ad* code for teacher A compared to the description for this same code for teacher B.

The interview data was collected and coded in the same way in order to gather evidence to answer the first research question as well. The online teacher log coded descriptions along with the interview response coded descriptions provides evidence of differences in the enactment of formative assessment by these two teachers. Again, for the interviews of the case study teachers the same coding system was employed.

As the teachers answered questions focused on Reflective Assessment their responses that were reflective of any of the four steps of RA were recorded next to the corresponding code on the data chart for that interview. There were six interviews for each teacher, thus six data charts of coded steps of RA descriptions were collected from each teacher for comparison. This, again, allowed for comparison between each teacher of their descriptions of RA, step by step.

The video data was collected and coded in the same manner as the online teacher logs and interview responses, but with a minor difference. The teacher actions were observed and recorded from the beginning of the lesson to the end, with a time recorded each time the action changed. A code was then given if the action was descriptive of one of the four codes of RA. To be clear, the online teacher logs were designed in such a way that teachers are allowed to express perceptions of RA starting with step one through step four. The coding of these expressions follows the data chart in this order accordingly.

The same concept holds true for the interview responses in that the questions guiding the responses are organized from step one through step four. This, too, allows for coding in sequential order of the steps. The video data, on the other hand, comes from recorded lessons from the teachers therefore observing and recording teacher actions must occur before coding of these steps is recorded. In essence, not every action observed in the video recordings were representative of a RA step, and some videos only provided evidence of enactment of one or two steps of RA that could be coded.

Data Analysis

The collection and coding described to this point provides an explanation of how the data collected from the online teacher logs, interviews, and video recordings was obtained and organized. Coding the data according to the steps of Reflective Assessment (a formative assessment strategy) for each of the case study teachers was meaningful mainly because it provides the qualitative evidence of formative assessment use. After the data was coded, I isolated the data coded for each code. I then separated it by research question one and research question two. I then wrote narrative summaries for each data source. The following paragraph describes this more specifically.

Qualitative Analysis of Online Teacher Logs and Interviews

Each teacher, as previously stated, completed three online teacher logs after a given Earth science lesson on water. Each online teacher log completed by the two case study teachers (appendix A) was completed after a lesson in which they focused on a particular concept relating to water. Analysis of the coded data, then, involved taking the data chart for the corresponding log and making a comparison of each step for the case study teachers. For instance, a comparison of the coded data from log one from teacher

A and likewise from teacher B was made and the outcome of this comparison was then noted and recorded. This process was repeated for each subsequent log.

The six interviews, five after lessons and one final, provided data that was collected and coded as previously described. A comparison of this coded data from each interview of each teacher was made in the same manner as the online teacher logs and the yielded results were recorded.

Qualitative Analysis of Video Recorded Data

Video recorded lessons provided data that was described and coded according the steps of RA as well. The recorded descriptions/codes and times for each video for each teacher were compared when possible first. For instance, video three for each teacher was of the same module lesson, so the comparison of the recorded descriptions/codes and times were made and results were recorded. For lessons that were different the same comparisons were made, but comparisons were not as specific. To be clear, video three was of the same module lesson, so specific lesson components were present in both teachers' videos. In videos where the lessons were not of the same module lesson for each teacher, the comparisons of teacher actions/codes were more general in nature.

CHAPTER IV
RESULTS

The research question this study set out to answer is reliant on the data collected and qualitatively analyzed of the two case study teachers. To be clear, the data that was coded according to the particular step of reflective assessment and expressed verbally through interview responses of the teacher, written by the teacher on the online log, or represented in practice by the teacher on video, was used in this study for purposes of answering the two research questions. Zeroing in on the four steps of reflective assessment that each teacher described and engaged in, as collected and coded from the data sources, yielded interesting results.

The interview data and the online teacher log data that was analyzed according to the four steps of reflective assessment is the qualitative evidence that attempts to addresses the question of:

> How do elementary teachers think about and engage in formative assessment differently to support students' learning of Earth science concepts?

Collecting and analyzing evidence of how teachers think and engage in actions related to formative assessment can provide insight for further study on the implementation of formative assessment strategies by elementary teachers. Identifying the differences that exist among elementary teachers' perceptions and resulting classroom actions related to formative assessment is a critical first step for improving the use of this type of assessment in elementary science. Through the online teacher logs and interview responses it is clear that the teachers enacted all four steps of RA. It is also apparent that differences in the teachers' perceptions and enactments of these steps exist. It is these identified differences that are critical in addressing research question one.

The online teacher logs submitted and interview responses given from each of the case study teachers were, for the most part, centered on the same lessons given by the teachers. For example, online teacher log one was submitted for the same Earth science water unit concept lesson for both teachers. This was the case for most of the online teacher logs and the interviews conducted. This is important because the differences in RA step perceptions and enactment descriptions of two teachers implementing the same lessons gives valuable insight into formative assessment enactment that has historically received little attention. Ultimately identifying how teachers use formative assessment differently in giving the same science lesson can provide a starting point for improving the use of this type of assessment.

Results from Log and Interview Data Analysis

There were two key findings from the analysis of the log and interview data; 1) the two teachers engaged in the first three steps of RA similarly, and the choice of next-steps actions (the *adjust* step) differed more than the other three steps of RA, and 2) the rationales of the two case study teachers behind decision-making in differed greatly. These findings are explained in the next subsections.

RA Step Engagement Findings

The first clear finding from the data analysis is that the teachers describe *anticipate, review, and reflect* steps they enacted, with minor differences in expression. In essence, both teachers; identify the concept of focus and student work used to gather student evidence (An), make a quick count of students who "get it" and those who "still need help" or (Re), and identify misconceptions from the student evidence (Rf).

For example, the surface tension of water was the Earth science concept that teacher A indicated log one was based on. This is indicative of the *anticipate* step of RA. Though teacher A identified the surface tension of water as being the concept she was focused on for this lesson, she did not indicate the student work that she would be using to examine for student understanding. Teacher A did identify 21 students that "still need help" understanding the concept in the *review* step of this lesson. Though the teacher did not give the student work she used to collect this number, the teacher does explain this number on the online teacher log therefore it can be assumed that some sort of student work was used in the *anticipate* step. The teacher wrote that her students "*did not get it"* and went on to explain that despite the concept focus of absorption of water, 13 out of 22 students *"had misconceptions."* This is an idea that is categorized as a *reflection* as the teacher is addressing what she noticed by reviewing student work.

Though teacher B completed the log for the same Earth science unit topic of water and also focused on the same concept of surface tension, the descriptions given for each of the four steps differ. In the *anticipate* step, teacher B indicates that surface tension is indeed the concept, just as teacher A did, but teacher B explains that she is *"looking for words stating that water can hold together forming a dome shape at the top."* Teacher B makes no mention of the work she will use to examine student understanding of the concept of surface tension, though in the *reflect* step she reports that five students *"got it"* and that 16 students *"still need help"* which implies that there actually was some sort of student work examined by this teacher.

The *anticipate* and *review* steps have minor explanatory differences, but are mostly describing the same main ideas of these corresponding steps. In the same regard, the teachers are both identifying misconceptions noticed in the *review* step as well. On the other hand, teacher A only briefly states that the surface tension of water is the focus of the lesson in the anticipate step. Teacher B gives an explicit explanation of what she is specifically looking for. Teacher B also identifies the specific misconceptions that students are giving in their responses. She also identifies that students are giving incomplete answers and that there is a definite need for further instruction. In contrast, Teacher A merely states that her students did not "get it" and "had misconceptions" in the *reflect* step. Though a minor detail, after comparing these responses in this first log it is clear that the detail that teacher B is providing in her logs is much greater than teacher A.

In the second online teacher log both teachers indicate the concept of focus, but teacher A expresses the student work she would look at and specific evidence she was looking for. Log two for teacher A is the Earth science topic of water, as expected, but with a focus on the concept that water expands as it freezes. The teacher explains that "I check lesson 2, part 3" is the student work she used to identify student understanding. The evidence that she expressed that she was looking for was that *"water expands when freezing."* By addressing the concept of focus for the lesson and the student work used for examination, teacher A successfully expresses defining components of the *anticipate* stage in this log.

The *review* step, as previously mentioned, deals primarily with identifying the number of students who understand the concept compared to those who do not. Teacher A entered 11 students in the *"got it"* category based on the indicated "I check lesson 2, part 3" student work reviewed, and seven students who *"still need help."*

Based on reviewing the student work, the teacher then was able to see that a misconception might exist about the expansion of water when freezing for her students. Teacher A explains in the *reflect* step that some students believe that the bottle got cold or froze, as opposed to the water inside the bottle, causing the cracking that was shown.

The Earth science topic of water with a focus on the concept that cold water contracts/warm water takes up more space was identified as what teacher B was focusing on in her second log, which meets the *anticipate* step definition. At the same time, though, she did not indicate the student work that she would be using to determine understanding of this topic by her students. This is an important aspect of the anticipate stage that teacher A left out. Student work *was* reviewed, though, as the teacher tallied 12 students in the *"got it"* category and nine in the *"still need help"* category, though it is not certain the type of student work that provided these results.

In the *reflection* step, teacher A indicated that *"troubles with reasoning"* was a trend she noticed after reviewing the student work, though she did go on to explain that all of the students did understand that the *"water level of the straw will go down when the bottle is cooled."* Teacher B left out the student work and specific evidence she was looking for.It is clear that the teachers engage in the *anticipate, review* and *reflect* steps of RA in relatively similar ways in looking at the results of their online teacher log data analysis.

The interview data analysis yielded further insight into the teachers' enactments of the first three steps of RA. Though there are differences identified in certain minor aspects of each step, each teacher is nonetheless engaging in first three these steps in similar ways as expected. The analysis of the interview data supports the online teacher log finding that the teachers do describe all four steps of RA, and that the first three steps of RA are expressed similarly.

The identification of the expression of perceptions or rationale relating to the steps of reflective assessment was coded for each step of Reflective Assessment (*anticipate, review, reflect,* and *adjust*) for each interview of each teacher. This allows for an analysis of the perceptions the teachers hold about Reflective Assessment. This data also gives insight into the steps of Reflective Assessment the teacher was/was not engaged in for the given lesson the interview was based on. This allows us to see how the teachers implement the intended four steps of the reflective assessment strategy and their thinking behind decision-making within each step.

Perceptions the teachers hold about Reflective Assessment can be identified in the final interviews as well. Each teacher was also interviewed a total of six times; five times after teaching a given lesson and a final interview at the end of the year. Again, each step of reflective assessment was identified in the transcribed interview when the teacher expressed the corresponding characteristics of the step. For example, if the teacher was explaining how she collected students' ideas about a concept of focus for a lesson this was coded as "An" to indicate that the teacher was describing actions that fit the *anticipate* step of reflective assessment.

For example, in Interview Two the concept of surface tension as it relates to water was identified and described by teacher A, or the *anticipate* step. Her description identifies surface tension as the concept and then she puts the concept into her own words. The following is teacher A's description of the concept:

> I would hope that surface tension, how water formed, it sticks together whenever you drop it on something else or see a drop of it, why it looks the way it does. (Interview Two, teacher A)

The lesson of focus for the second interview of teacher B was different than the lesson teacher A's interview was based on. Teacher B's lesson for this interview was focused on how water behaves on a slope and how the size of a water bead affects its ability to run down a slope. Teacher B identifies these concepts as the key concepts of this lesson and explains that she used an I-check from the FOSS kit as the student work she used to see evidence of student learning. In addition to these *anticipate* characteristics of Reflective Assessment, the teacher explains:

> What I did is I looked at, did they use the words "surface tension" and then water can connect, hold itself together. Those are some things I was really looking for, and then looking at their picture. (Interview Two, teacher B)

For each of the interviews the teachers describe these first three steps of RA in response to the corresponding interview question. This, again, was found to be relatively similar for both teachers in that they gave responses that were clearly indicative of the step of question. These teachers were essentially engaging in the first three steps of RA in the same manner as seen from the online teacher log results.

Important differences, though, are seen in the teachers' choice of next-step actions of the *adjust* step, more specifically the explanations provided for these choices. This is the second major finding from the data analysis in answering research question one.

Both the online teacher log data and interview response data support the claim that these two teachers' choices of and explanations for next-steps actions were very different.

In log one teacher A chooses the next step actions of "act" being a drop of water and "sentence starters." Teacher B does indicate the use of "sentence starters" as a next-step action, but also selects "class debate, re-writing, and modeling with a demonstration" as actions. She also describes her rationale for these choices in addressing student misconceptions. The teacher planned to use sentence starters as a next-step action to spark discussion by her students. Teacher B indicated that *"students are mixing up the water cycle and surface tension"* and that they are not providing enough detail in their description of their knowledge as her reasoning for choosing her next-steps actions.

Another difference identified in this log was teacher B's insight provided in the *reflect* and *adjust* steps. Teacher B specifically identifies that students are "*troubles with reasoning*" and accordingly chose next steps actions of "experimenting" with hot and cold water and a discussion of key points with the whole class. She explained her rationale for these choices as well. Teacher A, in contrast, chooses "revision with color" as a next-step action when she identifies the student confusion. The clear difference here is the rationale she provides for this choice. Teacher A's rationale for these actions is centered on improving student focus to detail. She also explains that this next step action was something she'd never done before and she wanted to try it out.

For addressing the misconceptions identified in log two, teacher A expresses that the next step actions of *"revision with color"* (in which students use different colors to indicate correct answers, incorrect answers, and added new answers) might *"improve focus to detail when explaining thinking."* She also wrote that she hadn't tried this particular next-step action which leads her to "*give it a shot*." In the same log, teacher B recorded that her next-step strategies would include experiments with hot and cold water and a class discussion of key points. The rationale she provided for these strategies was that they would *"reinforce concepts of expansion and contraction."*

In the third log, misconceptions identified in the reflect step lead teacher A to plan a next-step action of *"revision with color."* The teacher reported that this action allows students to confirm, cross-out, and add to their previous answers. Teacher A also reported that she needed more practice with this strategy. For teacher B next-step strategies of key points, which involves a whole class discussion of the *"key points"* of the concept; *"revision with color,"* which involves making specific types of corrections, additions, and deletions in the students' journals; and demonstrating the concept by freezing a bottle of air and a bottle of water were selected. Though the "revision with color" action is chosen by both teachers, the "key points" discussion is also chosen by teacher B in addition to this. From the analysis of this coded log data it may seem that the difference in choices of next steps actions by each teacher are minor, but the interview response analysis provided further insight into these log expressions. What I mean by this is that though each teacher may have chose "revision with color" as a next step action, their explanation for this choice is further elaborated in their interview responses.

The analyses of the interview data, as referred to earlier, yielded differences in *adjust* step choices. For example, in interview two of teacher A she describes the factors that she is going to take into account in planning the next lesson. She explains that she feels that having the students draw while giving their written explanation will help to "ingrain" the concept of surface tension a little more. She also expresses her belief that kids don't take the time to read and that visual actions might be more appropriate.

> I just think kids just don't want to take the time to stop and read stuff, sometimes the simpler the better. They can see the eyedropper, they can see 15 drops, and they could see the penny. I don't know if you've ever done it, but maybe work a try maybe? (Interview Two, teacher A)

Interview two of teacher B was focused on the same module lesson. The next-step actions that were taken to address misconceptions by teacher B involved using groups that included "experts." The teacher felt it was a good for students to be able to go to another expert student to talk to about their misconceptions. The teacher identified this as a *class debate*.

> The one thing that when you go into smaller groups, you have a greater chance of more interaction versus the large group. I wanted to make sure…go small group and then come back as a large group, and then show the demo what was going on. (Interview Two, teacher B)

This rationale behind the next-step action taken, or the adjust step, demonstrates a clear difference in thinking compared to teacher A's rationale for her next-step action taken. The idea of having students "interact" compared to the idea of "ingraining" a concept in students. In each of the five interviews of the teachers, differences in *adjust* step decisions emerged when comparing the two teachers' responses. In some instances the same action might have been chosen, or collaboratively decided, but the explanation for making this decision was much different. What is clear is that this last step of RA, in which the teacher makes a decision for what to do next to address the misconceptions identified, ultimately is most heavily reliant upon the teacher's discretion. The rationales described for this decision-making lead to the second major finding in answering research question one.

Perceptions and Rationales

The second major finding in answering research question one was that the reasoning being decision-making for each teacher differed greatly. Imperative in this finding is a difference in the teachers' perceptions of students and learning. This finding is significant because teacher A's rationale is reflective of a traditional or behavioral approach to teaching, whereas teacher B's next step choices and the reasoning behind them reflect a more constructivist approach to teaching.

For clarity, the intent of this study was not to specifically identify and categorize each teacher as a behaviorist or a constructivist teacher. Identifying differences in perceptions and enactments of RA between these teachers to support student learning, though, does necessitate a means of describing these differences. For the purposes of describing differences observed between teacher A and teacher B, a spectrum of behaviorist characteristics on one end and constructivist characteristics on the opposite end is used.

Through the choices of next-step actions in the *adjust* step of RA, and, more critically, the corresponding reasoning expressed by the teachers, it becomes clearer where the differences exist between these two teachers in implementing the steps of RA in their classrooms. The online teacher logs are only one means of gaining insight into the differences that exist between these two teachers' perceptions and enactments of RA. Analysis of the interview data is intended to provide a larger body of evidence to gain a clearer understanding of these differences. The contrasts observed in analyzing the logs of these two teachers, in the very least, suggest that they have differing perceptions on how students learn. Teacher A's log analysis suggest her view of student learning aligns more with a behavioral approach to teaching whereas teacher B's analysis aligns more towards constructivism.

A summary of each of the teachers' perceptions and ideas about RA is given below based on the online teacher log and, more significantly, the interview data. The final interviews of the teachers were based on a protocol of the following questions:

a. How would you define reflective assessment in the elementary science classroom?
b. What are the strengths of using RA?
c. What are the challenges of using RA?
d. How much did you use RA before this project?
e. How many times a week (on average) do you use (or do you plan to use) any part of the RA cycle?
f. How often do you go (or hope to go) all the way through the RA cycle?
g. What have you learned (or do you hope to learn) about your students as a result of using RA in your classroom?
h. Which next-step strategies are you most comfortable with?

These questions very clearly established the opportunity for each teacher to describe their perceptions of RA, learning and their students. In lieu of making the claim that teacher A's perception of student learning is, in fact, behavioral, I will instead only say that teacher A's *expressions* are behaviorist in nature compared to teacher B. Again, it is beyond the scope of this study to tease out the specific pedagogical characteristics of each teacher and classify each teacher according to a learning theory. It is, though, meaningful to note that there are distinct differences in the reasoning provided by each teacher that indicates their view of how students learn. When I describe characteristics as being behavioral in nature, or reflecting behaviorist thinking it should not be implied as classification of the teacher, but instead a mere description of a difference in reasoning that was observed.

Teacher A

Teacher A's online teacher log analysis indicated that her submissions are representative of a behavioral viewpoint of student learning. This claim is based on the analysis of three online teacher logs completed by teacher A. Further data on Teacher A's viewpoint, though, emerges from analyzing her interview responses and supports this claim.

It is assumed that Teacher A's perceptions of her students and how they learn ultimately, in some way, affects her perceptions and use of the individual steps of RA. Accordingly, what was found was that Teacher A's perceptions of student learning, as expressed in her interview responses, aligns more with the traditional, or behavioral end of the spectrum of learning. This was evidenced by some of the responses Teacher A gave relating to the steps of RA.

In the first interview with teacher A, she explains her concern about if students are capable of remembering what they've learned in the lesson two hours after the lesson had ended. This was given as her rationale for when she chose to assess student understanding from the previous class meeting's lesson. She goes on to explain that she feels students need a lot of engagement in terms of *how often* they are exposed to it, as opposed to *how* they are engaged with it. Student learning and retention of knowledge through repetition *(how often)* aligns with behavioral theory on learning whereas a focus on engaging students in meaningful way *(how)* is characteristic of a more constructivist approach to learning

In this same lesson, teacher A reviewed the student work and reflected that almost all of the students didn't understand the concept of focus. Based on this reflection she chose the next step actions of "sentence starters and acting out" and indicated that she needed to *"start from scratch."* In this interview she focused on recall, or retention of information, by the students to explain her reasoning for these choices.

As a whole, these responses from teacher A may give insight to her perceptions about student learning, but more importantly represent a clear difference from teacher B's perceptions about student learning. Teacher A conveys that being able to demonstrate understanding of a concept after a period of time has passed is important in determining if students understand the concept. This means that *retention of answers* is an important aspect of student learning for teacher A. These responses also imply that *repetition* and *retention* are important aspects teacher A considered when implementing the steps of RA in her classroom and thus establishes the clear difference between Teacher A's view and implementation of RA compared to Teacher B.

This claim is made under the assumption that teacher perception of how students learn influences teacher decision-making in their educational practices, such as RA. It can be said here, then, that Teacher A's expressions of behaviorist-like thinking impacted her implementation of the steps of RA. Her rationale and justifications relating to the steps of RA, more specifically *anticipate* and *adjust* steps, are expressions of behaviorist-list characteristics. When teacher A refers to retention, repetition, and the ability of students to recall information she is conveying the idea that students are perceived as passive receivers of knowledge provided by the teacher. It should be stated that this is in no way a judgment of good or bad for either case study teacher, but merely a description of clear differences that were observed.

The final interview of teacher A was conducted in an attempt to understand her views of Reflective Assessment at the end of a school year. Teacher A's understandings of RA, her students, and the decision-making rationale for next-step strategies are expressed and clear differences in these understands compared to teacher B's are observed. The definition that teacher A gives for RA is consistent with the expressions that she gave in both the online teacher logs and post-lesson interviews. This definition describes RA as a quick look at a concept on a regular basis allowing her to continually adjust her instruction.

This, again, focuses on a more behaviorist viewpoint in that she refers to the number of RA incidents and she focuses on her ability to adjust instruction. This isn't, by definition, constructivist, but it does identify that teacher A defines RA from an instructional viewpoint. She points to the value of RA as giving a more accurate account of what students actually know and expresses her disappointment that before her use of RA she assumed student understanding. Also, teacher A felt that a challenge of RA's implementation was that it was much slower and didn't allow her to get to lessons that she normally did get to without its use.

> Just finding that class time and just making it and we have. We've made it important enough to do it, but then unfortunately then, we don't get to some of those activities towards the end of the (inaudible). (Final Interview, Teacher A)

Teaching for coverage of topics as opposed to depth of understanding is descriptive of traditional science instruction rooted in behavioral learning theory. Teacher A's mention of this, then, further supports the differences identified up this point. Teacher A does note that RA's slower pace is beneficial in that it gives:

> …a magnified view of each little step along the way, so we really know if they're getting it or not, rather than waiting until the end. (Final Interview, teacher A)

The last clear difference identified in the final interview responses of teacher A relates to her perception of how long one full RA cycle takes to implement. Two class periods is given by teacher A as the length of time for one full cycle of RA. Along with this, her most comfortable choices for next-steps were "line of learning" and "revision with color." I will later address these choices of next-step actions compared to teacher B after I have summarized teacher B's final interview responses.

Teacher B

The online teacher log analysis of teacher B's expressions ultimately lead to a claim that teacher B's perceptions and enactments favor a constructivist viewpoint of teaching and learning. To further explore this claim, teacher B's interview responses were analyzed. Analysis of teacher B's interview responses further support this claim, and it is teacher B's perceptions and ideas that are in sharp contrast to Teacher A's. Another thing that becomes apparent while analyzing the interview data is that Teacher B's responses go into more detail with clear expressions of the factors guiding her decision-making for her lessons.

Based on Teacher B's interview responses, her view of student learning is constructivist in nature. It is assumed that Teacher B's perceptions of her students had an influence on the decisions she made about how to plan her lessons. In the first interview with teacher B she gives an explanation of her perception of her students compared to students in recent history. This is given as a rationale for the *anticipate* step in which she decided to give a formative assessment at the end of a lesson in an attempt to eliminate issues she's experienced with students in the past. The identification by teacher B of the nature of her students and the use of this knowledge in guiding her lesson planning is indicative of constructivist thinking. To provide clarity, the following was teacher B's statement regarding anticipated misconceptions and the nature of her students:

> In the past some of those things that always come up. I tried to prevent from happening just to make it correct so that they can kind of see it. Overall, I really thought it worked out pretty well…
> Today I think they're just wanting a quick answer. It's different today than five, ten years ago when I was teaching. I think because we're looking at a generation where they can go on Google. They're seeing their parents doing just that. Give me a quick answer. I'm not worried about why, I just want the answer. (Interview Two, teacher B)

Teacher B explained that her lesson was focused on students experiencing water under different conditions in order for them to make inferences based on their observations. She later justifies her rationale for enacting her lesson this way as a good way for her to circulate and check for understanding. Again, this description aligns with characteristics of constructivist theory on teaching and learning.

In her final interview, the definition teacher B gives focuses on RA as being a guide for what she needs to do next with her students. Teacher B refers to the value of RA in informing her goal-setting for students by providing her with a better account of student understanding. This description of RA from teacher B's perspective doesn't necessarily depict a constructivist viewpoint, but it is more student-centered compared to teacher A's definition of RA that highlighted frequency of RA use and instructional benefit, or a more teacher-centered depiction.

Another distinction expressed by teacher B in the final interview relates to the RA cycle frequency and length. Teacher A reported that two class periods was the amount of time she used to complete all four steps of RA. In contrast, teacher B explained that the amount of time needed was dependent on the individual lessons and how much in depth the lesson got. She went on to explain that she is using RA continually and across subject areas as well.

Both teacher A and teacher B felt that the greatest challenge RA posed was the amount of time needed to implement it. There are contrasting perspectives given, though, of how this increased time impacts their classrooms. On one hand, teacher A explained that she wasn't able to get to activities she normally gets to but that she sees value in the magnified view gained. On the other hand, though teacher B also reports not being able to get to as far as she used she goes on to explain that the time issue is diminishing as she embeds RA into her lessons. This represents teacher B's integration of RA as a part of her daily teaching strategies, which is an area that teacher A does not report doing.

Results for Video Data Analysis

The main differences that were found were how teachers enacted their lessons of next-steps strategies. More specifically the teachers' lesson enactment differed in discussion format, timing, and classroom management use. The video data that was analyzed provides an opportunity to view the teachers' actual classroom practices that are specific to the four steps of Reflective Assessment. Each of the recorded lessons was analyzed qualitatively in that descriptions were recorded of teacher actions throughout the lesson as well as the time in the lesson of the action. These actions were then identified according to the steps of RA. The identification of these steps of RA enacted by the two case study teachers allows for the description of what actual teacher implementation looks like for each step of RA. Ultimately, then, this should shed some light on the differences indentified the analysis of the online teacher logs and interview data

Coding this video by identifying the teachers' reflective assessment actions allows for a comparison of teacher A to teacher B with respect to differences in how they are engaging in reflective assessment with their students as well. This video data also serves to address the research question in regards to what the enactment looks observationally from the teachers.

It should be noted that the first three steps of Reflective Assessment (*anticipate*, *review*, and *reflect*) were identified and coded in the qualitative analysis of this video data, but the step of RA that was most prominent in the video data was the *adjust* step. The assumption can be made that differences that exist in the implementation of the first two steps aren't nearly as critical as the contrasting *adjust* enactments. The clearest differences lie in the decisions on next steps to address misconceptions, or the *adjust* step. In the next sections I will elaborate on the results of the qualitatively analyzed video recorded lessons of the two case study teachers.

The third video recorded lesson most clearly displayed the differences in Reflective Assessment implementation of the case study teachers, more specifically the next-steps actions enactments of the *adjust* step. For clarity, there were four video recorded lessons analyzed for each teacher.

The analysis of the video recorded lessons that were coded according to the steps of RA resulted in an overwhelming majority of *Ad* descriptions. This is mainly due to teacher A's submission of video recorded lessons that were of her implementation of next-steps actions. During the observation and analysis of the recorded lessons themes began to emerge for teacher A. As previously described, descriptions of actions were recorded throughout each lesson, from beginning to end, and appropriate codes of RA were identified when possible. What was clear after observing all four video recordings of teacher A was that her enactments of next-steps actions and module lessons looked very different than teacher B's.

Teacher A and B's times and actions are shown on the following figures:

Table 3: Lesson Three Time/Action Sequence for Teacher A

Time (min)	Action
0-3:00	"Acting out" particles of solids, liquids, and gases
3:15-4:18	Adding "volume" to water vocabulary list in journals
4:18 – 6:19	Discussion of density and predictions on vials
7:00 – 8:00	Students number off for partners
8:00 – 9:14	Students make vial observations
9:14 – 13:20	Students record and discuss observations as a group
13:20 – 16:00	Each group reports observations made to whole class
16:00 – 19:12	Whole class discussion of expansion/contraction of water and liquids versus solids
19:12 – 22:58	Blue ice/green water observations
22:58 – 24:45	Whole class discussion of "melting"
24:45 – 27:45	Students record observations
27:15 – 30:10	Whole class discussion of "What happens to water when you freeze it?"
30:00 – 32:20	Thermometer measurements of water temp. at different levels
32:20 – 35:00	Discussion of temperature difference observed
35:00 – 37:30	Discussion of sinking ice cubes
37:30 – 39:40	Clean-up and further syringe/vial observations
39:40 – End	Journal and exit slip writing: "Why do water pipes break when they freeze?"

Table 4: Lesson Three Time/Action Sequence for Teacher B

Time (min)	Action
0 – 10:00	Discussion of prior class period and syringe and vial of frozen water
10:00 – 11:15	Teacher hands out syringes/vials and discusses how to analyze syringes safely and appropriately
11:15 – 14:04	Students make observations as teacher circulates group to group
14:04 – 20:00	Whole class discussion of observations made/volumes recorded
20:30 – 26:09	Teacher demonstrates blue ice/green water procedure and hands out materials
26:09 – 29:00	Whole class discussion of observed results
29:00 – 30:32	Discussion of recess fountain being turned off during winter
30:32 – 33:40	Teacher demonstration of temperature measurement
33:40 – 36:00	Video shown and discussion follows of frozen pipes bursting.
36:00 – 42:00	Discussion of day's results of measurements/journal question writing/student sharing of journal writing
42:00 – 43:30	Students stand push chairs in if their responses matched correct response shared by a student
43:30 – 45:00	Students "act out" particle movement of solids, liquids, and gases.

To provide clarity, I will describe the narrative summary of video three for teacher A. At the beginning of this lesson teacher A's students are "acting out" what a solid, liquid and gas look like in terms of their particle movement. This is the next-step action identified in the online teacher logs and interview responses as "acting out." Selecting a next-step action based on student misconceptions identified in the reflect step of formative assessment is the essential element of the *adjust* step of RA. Teacher A guides students through this action by simply asking the whole class to show her what a solid, liquid and gas look like, respectively, in terms of their particle movement. This action accounts for the first three minutes of the lesson and is followed by the students sitting back down at their desks and pulling out their "water vocabulary word bank" at the request of the teacher.

The next portion of the lesson progresses from a whole class discussion about the prior lesson on density to a whole class discussion on ice and water. During this time the students observe frozen water vials report out their group findings. This eventually results in the teacher providing the explanation of the frozen vials looking different due to the water expanding when frozen.

Approximately twenty minutes into the lesson the teacher begins to discuss the experiment that the students will be performing. This begins with a whole class discussion of blue ice that was placed into green water. Teacher A asks the class what is happening to the blue ice and there were several answers given at the same time, such as "dissolving" and "swelling." Students were then instructed to record their observations in their journals.

Approximately 27 minutes into the lesson teacher A asks, "What happens to the volume of water when you freeze it?" A discussion ensues about last week's actions. This leads to a thermometer measurement of blue/green water where students then report back their measurements. The teacher then asks, "Why does the temperature drop?"

The lesson continues with whole class discussion of temperature differences at the top and bottom of the cup. Then there is yet another whole class discussion of the syringes that were observed at the beginning of the hour. The class ends with the class writing in their journal and on an exit slip about why water pipes break when they freeze.

Video three for teacher B is of the same lesson enactment that teacher A recorded. Teacher B begins the lesson by conduction a whole class, teacher-led discussion of frozen water in a syringe. She asks the students to hypothesize what the syringes are going to like, one filled with water, and one empty.

At ten minutes into the lesson, teacher B takes the frozen vials out of the freezer for students to observe. Students wait quietly to receive their vials for observation and then being to analyze them after a brief reminder from teacher B of safe handling. Teacher B then circulates from table to table of students addressing questions from students.

Four minutes later teacher B discusses the observations made with students through a question and answer session after which students record their measured volumes in their journals.

At twenty minutes into the lesson, teacher B begins to demonstrate the green water and blue ice interaction and then allows students to make observations on their own. As students are making observations, teacher B circulates to each group to provide the necessary materials. Ten minutes later a whole-class; teacher-led discussion of observations is conducted. The teacher then demonstrates how to make a temperature measurement and students make measurements and record according.

Teacher B then shows a brief video of a demonstration of water properties of expanding when freezing. This leads to a discussion of frozen pipes bursting as heard about in the news by students and students then write in their journals.

Something interesting to note is that both of the teachers filled an online teacher log before enacting this video recorded lesson. Both teacher A and teacher B indicated in their online teacher log that they would use “revision with color” as a next-step strategy for this lesson, which was not observed in the video analysis. Teacher A only indicates this as a next-step action in the online teacher log that she completed, though she was observed enacting discussions of key points during the video analysis. In addition to “revision with color,” teacher B also indicates “key points” and “other” as next-steps actions she planned to implement in this lesson, and she is analyzed as doing so accordingly.

The third video recorded lesson, nonetheless, highlighted the differences in the enactment of the same lesson by the two case study teachers. Both of the teachers’ students observed frozen vials of water (one empty and one full), blue ice placed in green water, and discussed how frozen pipes break in this lesson. The teacher actions descriptions recorded for the *Ad* codes are different in three main ways; discussion format, timing, and classroom management use.

Discussion Format Differences

Teacher A and teacher B held teacher led discussions which consisted of the teacher asking a question pertaining to either a prior lesson or observations just made. Teacher A’s use of these discussions can be best described as frequent and informal. What I mean by this is that often the students would respond to the question immediately, without being called upon and more than one student speaking at the same time. Sometimes teacher A would acknowledge some of the response, but mostly she provided a clear explanation of the answer to the question.

Teacher B’s discussion format can be best described as systematic and thorough. Teacher B required students to raise their hand to provide input on questions posed, and a response to the input was given before moving on to the next student. The nature of teacher B’s discussions were summative in that they were used to summarize an event that just took place in class.

Timing Differences

The timing of teacher actions was the second difference identified. For example, teacher B's enactment of the "acting out" next-step action took place at the end of her lesson in the third video. This action took place at the beginning for teacher A. How often and when teacher led discussions differed for the teachers as well. Teacher B led three whole class discussions; at the beginning of the lesson in regards to the prior lesson, after the vile observations were made, and after the blue ice/green water observations were made. In contrast, teacher A leads whole class discussions on seven different occasions; a density discussion at the beginning of the lesson, before the vile observations, after the vile observations, before the water/ice observations, after the water/ice observations, a volume/freezing discussion, and a discussion of sinking ice cubes.

Not only are there more frequent, teacher-led whole class discussions by teacher A, her actions during the various observations of water properties by students are much different. While students are given the different examples of water properties to observe (frozen viles, blue ice cubes in green water), teacher A mostly stands at the front of the room looking around at the groups and occasionally moves from one side at the front of the room to the other side. Conversely, during this same portion of this lesson teacher B circulates to each desk and answers student questions and guides them towards the concept they are observing.

The utilization of the class period for this lesson is also a timing difference. The progression of this lesson is: a review of the prior lesson, vile/syringe observations, blue ice/green water observation, temperature/floating ice test, then applying the concepts to a real-life example of frozen water pipes breaking. Each teacher implements actions for each of the steps in the progression, but differ in the amount of time and in some cases the way in which the portion is taught.

As can be seen from the time/action charts, teacher A's use of discussions is more frequent, and the discussions are much shorter compared to teacher B's. Teacher B's actions include the use of demonstrations, a video of a demonstration, and a discussion of the recess water fountain to enact this lesson. Also, the length of time allowed for the common actions, such as making and recording observations, varies in length. For instance, students in teacher A's class are given just over a minute to make observations of their vials, whereas teacher B's class is given just over three minutes to make the same observations.

Classroom Management Differences

The third difference identified from the descriptions recorded and coded of the teacher videos is in the area of classroom management. When I refer to classroom management I am referring to the organization of the classroom, procedures followed by students, and the general way in which the allotted class time is used for this lesson. Teacher A's classroom is organized into desks of four or five pushed together into a group, much like teacher B's. Teacher B's students, though, raise their hand and wait to be called up to provide input, whereas teacher A's provide responses freely during discussions.

The procedural expectation differences refer to the students in teacher A's class being assigned a number, one through five, and each number is assigned a duty when different materials need to be collected and brought back to the group for observation. Teacher B, on the other hand, delivered materials to each group for observation. When teacher A is leading a whole class discussion, there is general conversation taking place amongst some of the students in the class. Teacher B is quick to wait for the attention of the whole class before she poses a question or makes an explanation. She is seen flicking the lights off, to which the students immediately get quiet. She is also observed as saying that she was waiting for listeners.

Summary of Results

In addressing research question one, this study suggests that how teachers engage in formative assessment differently to support students' learning of Earth science concepts is mainly attributed to their decisions on next-steps actions to address identified misconceptions. The two teachers' descriptions of their rationale, as described in the data coded according to RA step, fell along spectrum of behaviorist to constructivist learning theory, with teacher A's rationale aligning towards the behaviorist end and teacher B's aligning nearer the constructivist end. The results of this study also show that these teachers engaged in the first three steps of RA (anticipate, review, and reflect) very similarly.

Research question two focused on what the differences in formative assessment enactment looked like. Even when the teachers in this study employed the same next-steps RA strategy, what this enactment looked like was largely different. Three major differences in lesson enactment involving next-steps action were found in the areas of discussion actions, lesson timing, and classroom management when video recorded enactment was analyzed.

It seems that the thinking behind decision-making of next-steps actions holds the key in further explaining how teachers engage in formative assessment differently. Along with this, the actual actions taken by teachers in implementing those strategies determines what these differences look like.

CHAPTER V
DISCUSSION AND CONCLUSIONS

Discussion

This study was a comparative case study of two teachers and provides qualitatively analyzed data to describe differences in formative assessment enactment. This study gives qualitative insight into formative assessment enactment differences that can exist when elementary science teachers engage in the same science curricula with their students. What can be taken from this study is a greater understanding of formative assessment use by teachers that can be further researched to improve the effectiveness of its use. This study is founded on the assumption that formative assessment is critical for student learning because of the large body of research findings of constructivist epistemology. In essence, studies that provide greater understanding of formative assessment usage have merit in developing effective formative assessment enactment by science teachers.

This study also focuses on the actions of teachers in creating learning opportunities that research has shown is an essential component of student learning. Though Reflective Assessment provides a strategy for enacting formative assessment, differences can exist in this enactment of formative assessment in science classrooms. Teachers' actions when adopting and implementing assessment strategies like RA are important to observe in order to describe effective enactment of such strategies. For example, actions that are more successful might garner more attention when providing teacher education training or professional developments. All in all this study and studies like it provide descriptive insight into teachers' formative assessment actions that can be highly influential in improving this type of assessment for teachers.

Constructivism and Formative Assessment

This study brings to focus the critical role of formative assessment practices of teachers in the construction of knowledge for students. The founding educational theorists' (Dewey, Piaget, Montessori) ideas about how students learn clearly established that learners are not passive receptors of knowledge provided by the instructor as is indicative of behavioral learning epistemology (Ultanir, 2012). Assessment of what students currently understand about a given concept is based on learners' previous experiences and background knowledge. This, then, is important in building new knowledge. This, we know, has been established as an epistemology known as constructivism (Ultanir, 2012; Sandoval & Reiser, 2004; Cobb, Yackel & Wood, 1992). Successful enactment of formative assessment strategies, such as Reflective Assessment, has a fundamental role in the more cognitively oriented constructivist approach (Tennyson & Rasch, 1988) on how students learn. Research on constructivism indicates the complexity of student learning and explains students as active, social, and creative learners that develop cognitively through learning experiences (National Academies Press, 2000; Perkins, 1999). This is where the action of teachers becomes important to be able to describe and a compare in creating learning experiences. The nature of this study centered on the formative assessment actions of teachers of students in the learning environment and the formative assessment strategies the teachers were engaging in with their students is consistent with constructivist theory (Vygotsky, 1978).

What is interesting about this study is that though it is focused on a formative assessment strategy, known as Reflective Assessment, which is rooted in constructivism, the differences identified in the two teachers' enactments are essentially a microcosm of the behaviorist versus constructivist disparity that exists in science instruction today. Even though a constructivist-based strategy was being enacted, the rationales and perceptions of student learning that guided decision-making in this strategy still reflected behavioral views of students as receivers of knowledge, as was expressed and enacted by teacher A.

In fact, as indicated by prior research, elementary science rarely focuses on conceptual understanding of science but instead emphasizes student engagement and hands-on activities (e.g. Beyer & Davis, 2008; Eshach, 2003; Forbes, Biggers, & Zangori, 2013). This suggests that hands-on activities are valued by science teachers, but there is generally a lack of focus in conceptual understanding connecting to the hands-on activities. For instance, a teacher can set up and implement hands-on activities for students, but if there aren't plans in place by the teacher to identify and build upon student conceptual understanding, then the activities are meaningless as an educational endeavor.

This study is interesting in regards to this prior research. Engagement and hands-on activities were displayed in the videos and referred to in the log and interview data of the teachers in this study. It was shown that teacher A and teacher B enacted the same Earth science module lessons, but teacher A's focus was more on the retention of information of the students. She was also quick to provide explanations of concepts through frequent, whole class discussions. Teacher B, on the other hand, was more focused on the conceptual understandings of the students as evidenced by her more frequently circulating the room while asking thought provoking questions and her increased use of demonstrations and expressions of student understanding concerns.

For sake of clarity, despite both teachers providing the same hands-on opportunities in the lessons, the way in which the teachers addressed conceptual understanding of the students was much different. Clearly retention of teacher disseminated concept knowledge is a much different focus than the on-the-fly questioning and interaction with students during the hands-on activities. Activating student thinking during the hands-on activities connects the actions to the conceptual understanding and is constructivist in nature. Providing frequent explanations of concepts, by nature, assumes that students receiving knowledge from the teachers verbally and is behaviorist in nature.

Along these same lines, research has shown that teaching models, such as the FOSS curricula in this study, that are based on *engaging* students as an important first step in science learning and involve deciphering students' prior knowledge about the concepts of interest (Goldston et al., 2010; Wilson et al., 2010; Bybee et al., 2006). The RA strategy enacted during the implementation of the FOSS Earth science module by the teachers in this study is consistent with this research in that it specifically addresses assessing students' prior knowledge in the given module lessons, but ultimately it is the teachers' decision-making in this enactment that emerges as a key factor, especially concerning the next step actions after misconceptions have been identified. I will focus more clearly on decision-making in the next discussion point.

Teachers' Ideas and Decision-Making

The second discussion point that emerges from this study is the critical role that differences in teachers' ideas and decision-making play in formative assessment implementation. An implementation difference based on teacher decision-making asa evidenced from this study involved long-term lesson planning. Lesson planning involves close attention to sequencing the curricular units based on how the school year is designed for science instructional time. It follows that teachers' decision-making in this process ultimately determines the actual pacing of instruction as the school year progresses. What's interesting, though, is that research has indicated that formative assessment strategies have been shown to be easily implemented in elementary classrooms, taking little additional time for teachers (Osmundson, Dai, & Herman, 2011; Kennedy, Long, & Camins, 2009). I should point out that this doesn't describe the ways or types of formative assessment that is being implemented, which is what this study attempts to shed light on.

In this study, though, both teachers' indicated that implementing RA in the enactment of the Earth science modules didn't allow them to get as far in their science curriculum as they were used to getting. This might be reflective of the FOSS Earth science modules, and the use of RA in enacting these lessons, being more strongly rooted in constructivist epistemology. The act of collecting assessment-induced evidence of student learning and adapting instruction can also be described as assessment *for* learning, or formative assessment (Cauley and McMillan, 2010; Black and Wiliam, 1998b; Pellegrino, Chudowsky, & Glaser, 2001). If we speculate that collecting this evidence for the purpose of adapting instruction wasn't a critical element of the previous science instruction, then the time that is now given for this might slow pacing. In essence, it seems logical that engaging in RA while implementing the FOSS Water module in contrast to the prior sequencing and enacting of science lessons by the two teachers would be the source of the slower pacing.

As noted earlier, research on the benefits of formative assessment has pointed to the effectiveness in student learning gains (National Academies Press, 2007; Black and Wiliam, 1998b; White and Frederikson, 1998). What research hasn't shown are the teacher actions regarding formative assessment enactment and the potential impact on effectiveness their contrasting formative assessment actions might have. This study focused on identifying differences in formative assessment actions of two teachers enacting the same science curriculum in an attempt to gain a better understanding of the potential impact it might have. Benefits in terms of student motivation and thus increased engagement have ties to formative assessment techniques as well (Miller and Lavin, 2007; Boddy et al., 2003; Karlsson, 1996). This brings to focus the importance of studies of teacher actions, such as this one, in the implementation of formative assessment to begin to highlight specific teacher actions of greatest merit. Studies of the fidelity of implementation (Furtak et al., 2008; Pijanowski, 2008) have shown that how closely teachers adhere to prescribed formative assessment actions is linked to greater learning gains. Though this was not the central focus of this study, understanding that teachers' formative assessment actions can yield different learning results highlights the value of future research of teachers' formative assessment actions.

Numerous studies highlight the benefits of formative assessment in assessing for learning in a science classroom (Ayala et al., 2008; Furtak, 2008; Hermain and Choi, 2008; Heritage, 2007; Furtak, 2006; Ruiz-Primo & Furtak, 2006). The three distinct forms of formative assessment (Ruiz-Primo & Furtak, 2006) represent a spectrum of informal to formal assessments (Shavelson and SEAL, 2003) for learning of students. The FOSS curriculum the teachers were implementing in this study included *curriculum-embedded* formative assessments that the teachers used, so the prevalence of this form of formative assessment was essentially the same for both of the teachers.

The *planned-for* form of formative assessment is the form of formative assessment that relates to the on-purpose planning of the teacher in the given lesson. This type of formative assessment is interesting in that it is not prescribed by the curriculum, but is dependent on the pre-planning of the teacher. Where on-the-fly formative assessment is more reactionary, or in the moment, planned-for formative assessment requires foresight from the teacher to make decisions on when to enact this form. This, then, is an important form of formative assessment that warrants further study of the ways in which teachers engage in this on-purpose planning differently.

It was clear to see in this study that the teachers engaged in two of these three forms of formative assessment, with teacher B engaging in *on-the-fly* formative assessment during lessons more frequently than teacher A. Of the Reflective Assessment steps enacted by the teachers in this study, the *adjust* step displayed the largest number of differences between the two teachers. For instance, teacher B was shown to implement a greater diversity of strategies and asked more concept-eliciting questions. Research of *on-the-fly* formative assessment, or informal formative assessment, has shown that this form of formative assessment holds enormous potential for students and teachers in learning gains and teachers' perceptions of student learning and understanding (Gallagher, 2009; Crumrine & Demers, 2007; Heritage, 2007; Ruiz-Primo & Furtak, 2006). Future studies of the ways in which teachers address student understandings or misunderstandings on-the-fly might provide further insight into the dynamics of this type of teacher-student assessment action.

This study operates under the premise that reflective assessment strategies have shown benefits in recent research (Trauth-Nare and Buck, 2011; Kennedy, Long and Camins, 2009; Toth, Suthers and Lesgold, 2002). What this study attempts to address is the variability in the use of Reflective Assessment by teachers. The differences in the enactment of Reflective Assessment by the two teachers in this study mostly relate to their next-steps actions decisions and corresponding rationales once the teachers had identified themes or misconceptions held by their students. This difference might be due to the more behaviorist perceptions of student learning expressed by teacher A guiding her classroom action when enacting the FOSS lessons, though this warrants further study to make such claims of causation. Teacher B's differences identified from the video data compared to teacher A displays a greater use of on-the-fly formative assessment actions as well. The teacher actions related to formative assessment enactment, such as using RA, in implementing science curriculum undoubtedly involves many variables.

This study sought to gain a greater understanding of the use of formative assessment in science instruction by qualitatively studying two case study teachers' descriptions and actions related to the four steps of Reflective Assessment. The resulting differences found point to the need for further research into the factors that influence the enactment of formative assessment strategies by elementary science educators.

Limitations

This study is limited in that it looks at the formative assessment actions of two teachers in a comparative case study methodology which descriptive data that serves to provide greater understanding of this topic and does not address causation. Undoubtedly a comparative case study of two teachers in the classroom setting hardly provides substantial data to make sweeping claims, but it does at least provide results that can contribute to other similar research on elementary teachers' use of formative assessment. The scope of this study was to gain greater understanding of the differences in formative assessment use in developing student understanding in science, thus a logical next step might be to identify the source of these differences. Looking at the formal teacher training and other factors surrounding pedagogical content knowledge, for instance, might be considered in trying to pinpoint causation for specific formative assessment actions. It was beyond the scope of this study to consider the background information of the teachers beyond the superficial descriptions provided. This study did not address what the two teachers' prior enactment of science lessons looked like prior to RA training. Observing the teachers' enactment of science lessons prior to RA training could have possibly provided insight into the teachers' decision-making in regards to student understanding and the lesson/curriculum pacing. For example, how did the teachers identify student understanding and make decisions of next-steps actions prior to RA training and implementation?

At the same time, this study also was limited in that the data collected from the three sources only represented the first year of the implementation of both RA and the FOSS curricula. Taking into consideration how teachers interact with new curricular materials differently at the same time they are implementing the newly acquired RA formative assessment strategy might look different than implementing RA with their prior curricula. This might also look different in the second and third years after the teachers have had a chance to work out misunderstandings and other miscues. The FOSS module provided the materials for formatively assessing students, such as the "I checks" described by the teachers. These were used as the student work that provided the evidence of student understanding. Looking at how teachers might create or adapt different curricular materials in the implementation of RA steps might provide a clearer picture of differences in the engagement of the first three steps of RA that was found to be very similar while these teachers enacted the FOSS module lessons.

Overall this study serves to not only provide anecdotal descriptions of formative assessment use by elementary science teachers, but to potentially add to other similar research in gaining consensus on how to improve assessing for learning in elementary science.

Conclusions and Implications

A closer look at the differences in the implementation of formative assessment by elementary science teachers might expose important aspects for further research on the implementation of effective formative assessment. Studies on the use of formative assessment in enacting science curricula have great implications for teacher education programs for pre-service teachers and professional development for practicing teachers as well. If it can be shown that specific kinds of formative assessment enactment lead to greater learning gains for students, then the adoption and enactment of such actions should be vigorously encouraged by the stakeholders of science education.

Teacher education programs that provide the research-based framework of constructivism should then stress the critical nature of formative assessment in science learning. It should follow, then, that the formative assessment actions that hold the greatest potential for student learning gains ought to be the formative assessment actions that pre-service teachers learn to enact. It is important that pre-service teachers gain classroom experiences in enacting the more successful formative assessment actions if they are to have success in using these same actions in their future classrooms. Just like any other skill, good use of formative assessment entails practicing the successful skills. Identifying the successful components of formative assessment use, then, becomes paramount in developing good elementary science teachers. Teacher education programs that allow students to use formative assessment in classroom setting might help teachers to see the value of using this type of assessment in their future classroom. There is no doubt that having good elementary science teachers can have a great impact on the science learning opportunities of elementary students.

In-service teachers, on the other hand, such as the two teachers in this study, must have professional development opportunities to learn the research-based benefits of constructivist teaching and learning. The demands on elementary teachers to teach multiple subjects throughout their day, as opposed to secondary levels which likely teach science only, poses a challenge in achieving this. Though there is substantial research on the effectiveness of formative assessment for student learning, little attention has be paid to how elementary teachers use formative assessment in their class rooms. Nonetheless, formative assessment actions incorporated in the enactment of science curricula that lead to increased learning gains of elementary science students holds the key in raising student understanding of science.   The teachers in this study, for instance, both indicated that by using the RA steps to implement science lessons provided them an eye-opening experience. They were surprised by what their students didn't really understand and realized that their old way of teaching would not allow them to identify this. They would just move on after the unit test and assume students understood. This type of experience for other in-service teachers might ultimately hold the key in transitioning them towards using formative assessment by having this "a-ha" moment. This could be a critical gateway for teachers to transition to more cognitively appropriate teaching methods.

Comparative case studies of the differences in teacher enactment of formative assessment actions in enacting science curricula can provide increased understanding of the benefits of assessing for learning. How formative assessment is used differently can further elaborate specific actions and their effects. For instance, if differences can be categorized and evaluated, then the development of more specific formative assessment strategies is possible. Building a body of research that unmasks the beneficial actions of assessing students for learning can allow teachers to zero in on hone their actions when enacting formative assessment of their students.

Curricular materials can also benefit from a greater understanding of how teachers use formative assessment differently and what these differences look like. For instance, the FOSS modules make use of curriculum-embedded formative assessments in their materials to allow teachers to gather evidence of student understandings. Future development of curricular materials that take into account the research on how formative assessment is most effectively enacted allows for teachers to better enact rich learning opportunities for students. Ultimately the curricular materials that teachers interact with must be aligned with research-based effective strategies the teacher is enacting in the science lesson. Is should follow that curriculum developers take into account the research-based effective formative assessment actions in designing their curriculum for elementary teachers.

There are many benefits to researching effective strategies to raise student achievement in elementary science. Teachers bear a great burden of using effective strategies in an attempt to create learning opportunities that lead to greater understanding of science. The importance of solidifying effective formative assessment use for elementary science teachers cannot be underestimated in the efforts to increase effective, research-based science instruction.

REFERENCES

Ayala, Carlson C., Shavelson, Richard J., Ruiz-Primo, Maria Araceli, Brandon, Paul R., Yin, Yue, Furtak, Erin Marie, Young, Donald B., & Tomita, Miki K. (2008). From formal embedded assessments to reflective lessons: The development of formative assessment studies. Applied Measurement in Education, 21, 315-334.

Black and Wiliam (1998a) Black, P. & Wiliam, D. (1998-a). Assessment and classroom learning. *Assessment in Education* 5(1), 7-74.

Black, P. & Wiliam, D. (1998-b). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80*(2), 139-148.

Bybee, R. Taylor, J., Gardner, A., Van Scotter, P., Carlson Powell, J., Westbrook, A., and Landes, N. (2006). The BSCS 5E instructional model: Origins and effectiveness. A Report Prepared for the Office of Science Education, National Institutes of Health. Colorado Springs, CO.

Cauley, Kathleen M. & McMillan, J. (2010). Formative assessment techniques to support student motivation and achievement. *The Clearing House, 83,* 1-6.

Chiapetta, E.L. (2008). Historical Development of Teaching Science as Inquiry. In J. Luft, R. Bell, and J. Gess-Newsome (Eds.), Science as inquiry in the secondary setting (pp. 21-30). Arlington, VA; National Science Teachers Association.

Cobb, P., Yackel, E., & Wood, T. (1992). A constructivist alternative to the representational view of mind in mathematics education. *Journal for Research in Mathematics education*, 2-33.

Creswell, J. W. (2006). Qualitative Inquiry And Research Design: Choosing Among Five Approaches Author: John W. Creswell, Publisher: Sage Publica.

Daka, K. R., & Ballet, J. (2011). Children's education and home electrification: A case study in northwestern Madagascar. *Energy Policy, 39*(5), 2866-2874.

Donovan, M. S., Bransford, J. D., & Pellegrino, J. W. (2000). How people learn: Brain, mind, experience, and school: Expanded edition.

Furtak, Erin Marie. (2008). Formative assessment in k-8 science: A conceptual review. (Commissioned paper by the National Research Council for Science Learning K-8 consensus study.)

Heritage, M. (2007). Formative assessment: What do teachers need to know and do? *Phi Delta Kappan, October,* 140-145.

Kennedy, C., Long, K., & Camins, A. (2009). The reflective assessment technique. *Science and Children*, *47*(4), 50-53.

Lincoln, Y. S., & Guba, E. G. (1985). Effective evaluation. *New York: Jassey-Bass*.
Norman Kent Denzin, & Yvonna Sessions Lincoln (Eds.). (2005). *The Sage handbook of qualitative research*. Sage.

Oliver, G. (2004). Investigating information culture: a comparative case study research design and methods. *Archival Science*, *4*(3-4), 287-314.

Osmundson, E., Dai, Y., and Herman, J. (2011). *Year 3 ASK/FOSS efficacy study*. (CRESST Report 782). Los Angeles, CA: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).
Perkins, D. (1999). The many faces of constructivism. *Educational leadership*,*57*(3), 6-11.

Ruiz-Primo, M.A & Furtak, E. (2006). Informal formal assessment and scientific inquiry: Exploring teachers' practices and student learning. *Educational Assessment, 11*(3&4), 205-235

Ruiz-Primo, Maria Araceli. (2011). Informal formative assessment: The role of instructional dialogues in assessing students' learning. *Educational Assessment*, 37(1), 15-24.

Sandoval, W. A., & Reiser, B. J. (2004). Explanation-driven inquiry: Integrating conceptual and epistemic scaffolds for scientific inquiry. *Science Education*,*88*(3), 345-372.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference.

Tennyson, R. D., & Rasch, M. (1988). Linking cognitive learning theory to instructional prescriptions. *Instructional Science, 17*(4), 369-390. doi:10.1007/BF00056222

Terhart, E. (2003). Constructivism and teaching: a new paradigm in general didactics?. *Journal of curriculum studies*, *35*(1), 25-44.

Ultanir, E. (2012). An epistemological glance at the constructivist approach: Constructivist learning in dewey, piaget, and montessori. *Online Submission*

Wilson, C. D., Taylor, J., Kowalski, S. M., & Carlson, J. (2010). The relative effects and equity of inquiry-based and commonplace science teaching on students' knowledge, reasoning, and argumentation. Journal of Research in Science Teaching, 47(3), 276-301.

APPENDIX: ON-LINE TEACHER LOG

1. Which unit are you currently teaching?
2. What is the <u>key concept</u> you are assessing in the student work? What <u>student work</u> are you examining? What are you looking for in students' responses as <u>evidence of understanding</u> of the key concept?
3. After reviewing student work, how many of your students:
    a. Got it
    b. Still need help
4. What did you notice in the student work you reviewed? Did you observe any misconceptions? What trends in students' understanding of the key concept did you see?
5. Will you use a next-step strategy in your next lesson?
    a. Yes
    b. No

    If No: survey ends.
    If Yes:
    1. Which next step strategy (or strategies) will you use?
        a. Line of learning
        b. Group consensus/White boards
        c. Response log
        d. Class debate
        e. Key points
        f. Revision with color
        g. Review and critique anonymous student work
        h. Critical competitor
        i. Multiple-choice discussions
        j. Sentence starters
        k. Sticky note feedback
        l. Conferences
        m. Centers
        n. Reteach the concept with a small group
        o. Other

        If Other is selected:
            i. Please describe the 'other' next step strategy you used.
    2. <u>Why</u> did you select this next-step strategy? <u>How</u> do you think it will enhance students' understanding of the key concept?
    3. Do you plan to video record the lesson that includes this next-step strategy?
        a. Yes
        b. No

        If Yes:
        1. The number on the SD card I am submitting in the mail is _______.

        If No:
        1. Please explain why you chose NOT to use a next step strategy