A Dissertation

entitled

A Comparison of Two Classroom Assessment Contexts Following a Science Investigation: Does the Use of Interview as an Assessment Tool Provide Different Results than Existing Teacher-Driven Tests?

by

Tamala S. North

Submitted to the Graduate Faculty as partial fulfillment of the requirements for the Doctor of Philosophy Degree in Curriculum and Instruction: Early Childhood Education

____
Dr. Ruslan Slutsky, Committee Chair

____
Dr. Shauna Adams, Committee Member

____
Dr. Florian Feucht , Committee Member

____
Dr. Susanna Hapgood, Committee Member

____
Dr. Amanda Bryant-Friedrich, Dean
College of Graduate Studies

The University of Toledo

May 2018

An Abstract of

A Comparison of Two Classroom Assessment Contexts Following a Science Investigation: Does the Use of Interview as an Assessment Tool Provide Different Results than Existing Teacher-Driven Tests?

by

Tamala S. North

Submitted to the Graduate Faculty as partial fulfillment of the requirements for the Doctor of Philosophy Degree in Curriculum and Instruction: Early Childhood Education

The University of Toledo
May 2018

Assessment is often used to hold schools and teachers accountable for student learning. Assessment instruments are used as tools to guide curriculum choices and lesson plans, from districts to individual students. In any discipline, knowing what students know and what they have learned following a lesson is important on multiple levels. This is especially true in subjects like science. Despite recommendations from many leading early childhood education authorities, short answer, paper-and-pencil type, tests often prevail for young learners across many subject areas. This type of assessment may not provide all students with the opportunities needed to fully articulate what they've learned. This study examined two contexts of assessment following a science investigation, in a second-grade classroom - existing teacher-driven tests and an interview, or narrative, approach to assessment. The teacher-driven test was a one-dimensional, multiple-choice, test. The post-lesson interview was an open-ended interview where student participants were asked to tell a story about what they learned in science. During the interview, students were provided with an opportunity to use various other

“languages” to share what was learned. Although the teacher-driven test proved to be a successful assessment “language” for some students, the post-lesson interview empowered more students to express what they learned with more detail. The post-lesson interview also showed that this type of assessment context has the potential to provide useful feedback on lesson quality, future science investigations, and individual student interests. This study also applied an innovative approach to the assessment data by pairing the teacher-driven test *with* the post-lesson interview. Among other positive results, this unique pairing raised teacher-driven test scores from below average, to above average. This research demonstrated that one-dimensional assessments fail to capture everything students learn. Future research should focus on the creation and application of assessments that allow students to use their preferred assessment “language.”

Acknowledgements

I owe a debt of gratitude to my Committee Chair, Dr. Ruslan Slutsky. I would like to thank him for his patience, support, respect, and kindness during this journey. Thank you to Committee Members – Dr. Shauna Adams, Dr. Susanna Hapgood, and Dr. Florian Feucht. Their time and support through this process created such a positive experience for me. I am beyond grateful to them for serving on my committee. I would also like to acknowledge Dr. Amy Allen, who advised me well through my coursework and planted the seeds of thought for this dissertation.

This opportunity would not have been possible without the committee and staff of NURTURES. I'm grateful to Dr. Laurie Dinnebeil for opening that door. Thank you to Dr. Charlene Czerniak, for exemplary leadership in science education. And thank you, Deanna Bobak, for the support and resources needed to be successful. I am also grateful to my Early Childhood Education colleagues in the Judith Herb College of Education. Lilla Horton – thank you for everything you did for me. Dr. Jessica Branch, and Dr. Shelley McNally – thank you for sharing a major part of this journey with me and offering guidance and insight when needed.

Most importantly, I would like to thank my family for their confidence in me, and for the constant stream of encouragement. I'm especially thankful for my husband, Dr. Phillip North, who supported me in every way and during every step of this journey.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

C1................................Classroom One
C2................................Classroom Two

MPS............................Midwest Public School

NRC ....National Research Council

SP ....Student Participant
SP1C1 ....Student Participant (1-14) Classroom (1 or 2) Identification Code

Chapter 1

Introduction

Statement of the Problem

Classroom assessment has become a main focus in early childhood education, and an important component of teacher practice and professional development. Accountability and learning standards in early childhood education have been front and center and, as a result, assessment in early childhood education has been labeled a priority by both public agencies and private organizations (Snow, 2008). The National Association for the Education of Young Children's (NAEYC) current position on assessment in early childhood education states that assessment should be used to identify young students' progress, strengths, and needs. Methods used in the assessment process should be developmentally appropriate, culturally responsive, connected to students' daily activities, supported by professional development, and include the family. Additionally, the method of assessment must have a "specific and beneficial purpose."

Unfortunately, teachers do not receive sufficient assessment training prior to entering the classroom and often indicate they need additional training to become proficient in assessment (DeLuca & Klinger, 2010). Many teacher assessment practices focus on the process of answering a set of questions that offer a limited number of options for student responses (Enger & Yager, 2009). Howard Gardner (2011) warned that these tests, often administered by paper-and-pencil void of any interchange with the examiner, are part of the sad scenario that has repeated itself over and over throughout the history of education. He goes on to suggest that too much significance is often given to one single test score. And although many educators do not approve of such limited options to identify student progress, not much has changed in the pursuit to know what students know. Gardner dares us to believe that there is more to intelligence than short answers to short questions. Likewise, NAEYC stated in their curriculum, assessment, and program evaluation position statement that the results of single assessments are often unreliable since young students may not understand the importance of "doing their best" or may be negatively influenced by internal and/or external distractions (2003). Despite recommendations from these, and many other leading early childhood education authorities, traditional short answer summative tests often prevail for young learners.

Science assessment is particularly problematic in early childhood education, as the emphasis on science education is relatively new in the field. A smattering of research has made its way on to the scene, however, summative assessment methods with strong validity evidence allowing teachers to evaluate the impact of learning do not exist (Greenfield, 2015). The National Science Teachers Association (NSTA) suggests that the reason these methods do not exist is because they are challenging to design and require an extensive amount of professional development. Additionally, there is a need to track student progress in science education over the course of the academic year (formative assessment). Although some comprehensive assessments include a few science items, no formative assessment specifically designed for science education exists (Greenfield, 2015).

The current conditions surrounding assessment in early childhood science education should come as no surprise. Just a few years ago early childhood science education consisted of knowing the basic characteristics of living things, and classifying objects based on physical properties (Greenfield, Jirout, Dominguez, Maier, & Fuccillo, 2009). However, in more recent years some states have demonstrated a greater emphasis on science in the early years by designating the subject as a school readiness domain, making research in early childhood science assessment all the more important (Greenfield, 2011; Sackes, Trundle, & Flevares, 2009).

Purpose of the Study

The purpose of this study was to examine two contexts in which summative assessments tell us what students know. More specifically, this study documents what second-grade students can demonstrate to us related to their understanding of a scientific concept. Students were asked to share what they learned by engaging in two different contexts of assessment. The first context was the routine teacher-driven assessment method, and the second context employed an interview approach. The students were asked to tell the researcher a “story” about what they learned. The ability of each assessment to accommodate the student’s capability to share what they learned was studied.

For many years, early childhood education professionals have thought of science as nothing more than a vehicle for the development of young children’s social, physical, and language development (Worth, 2010). In more recent years, however, there has been a growing understanding that science education provides even very young children with an opportunity to engage in and develop critical thinking skills that will serve them in school and life (Harlen & Qualter, 2014). Additionally, there is a developing belief that science education is an important domain to be explored in the early childhood classroom (Worth, 2010).

In 2011, The National Research Council (NRC) and their partners embarked on a two-step process to develop the Next Generation Science Standards. The process was motivated by the belief that quality science education is based on standards that are rich in content and practice, consistent with aligned curricula, pedagogy, assessment, and teacher preparation. Prior to this time, it had been nearly 15 years since science education standards had been updated using research that provided us with a better understanding about how young children learn.

With more and more schools introducing high quality science experiences to their early childhood curriculum, so too comes the need for high quality assessments to monitor the success of the students, teachers, and the process of science inquiry. Upholding the philosophy of The National Science Teachers Association, teachers must recognize that appropriate measures need to be taken to ensure that all students receive the support and resources needed to be tested fairly in science education (2001). In science, like most genres of education, it is generally accepted that both formative and summative assessment have important roles in the classroom. One shared role of the two methods is to advance student learning. Though formative assessment plays the important role of scaffolding learning during the inquiry process, the achievement emphasis is often placed on the snapshot of success or failure that summative assessment provides teachers, parents, and administrators (Harlen & Qualter, 2009).

The context for summative assessment is especially important in early childhood settings because younger students typically have lower levels of communication and writing skills which could interfere with their ability to successfully express what they've learned. Accurate summative assessments are also important for teachers as they guide future instruction and provide feedback on the quality of past lessons.

This study focuses on allowing students to use “story” as an alternative summative assessment providing a way for students to communicate their level of understanding of scientific concepts taught in second grade (e.g., chapter tests, unit tests). It is concerned with the ability of existing teacher-driven summative assessments to provide an accurate representation of students’ conceptual understanding of a scientific concept. This research is not a comparison study involving formative and summative assessment. It is not an investigation of high stakes testing methods that govern student advancement. This study intends to better understand alternative contexts in which summative assessments can be executed, and discuss the results so that future efforts can inform teachers and provide students with the best chance to demonstrate, or verify, what they know.

Theoretical Framework

The Reggio Emilia Approach, as interpreted in *The Hundred Languages of Children* (Edwards, Gandini, & Forman, 2012), is the overarching theoretical framework for this study. This approach nurtures intellectual development through a systematic focus on symbolic representation. Learners are encouraged to make their thinking visible to others by using any of the “one hundred languages," or numerous ways, to accomplish this task. Within this approach to learning and teaching, children are encouraged to express themselves in their own unique ways and by using any “language.” The founder of the approach, Loris Malaguzzi, believed that young students are unique individuals who interact and learn in many ways. Malaguzzi wrote a poem to express this belief (Edwards, Gandini, & Forman, 2012):

No Way, The Hundred is There (translated by Lella Gandini)

The child
is made of one hundred.
The child has
a hundred languages
a hundred hands
a hundred thoughts
a hundred ways of thinking
of playing, of speaking.
A hundred always a hundred
ways of listening
of marveling, of loving
a hundred joys
for singing and understanding
a hundred worlds
to discover
a hundred worlds
to invent
a hundred worlds
to dream.
The child has
a hundred languages
(and a hundred hundred hundred more)
but they steal ninety-nine.
The school and the culture
separate the head from the body.
They tell the child:
to think without hands
to do without head
to listen and not to speak
to understand without joy
to love and to marvel
only at Easter and Christmas.
They tell the child:
to discover the world already there
and of the hundred
they steal ninety-nine.
They tell the child:
that work and play
reality and fantasy
science and imagination

sky and earth
reason and dream
are things
that do not belong together.

And thus they tell the child
that the hundred is not there.
The child says:
No way, the hundred is there.
-Loris Malaguzzi (1920-1994)

The Reggio Emilia Approach concentrates on every learner with respect and a willingness to allow them to express what they are thinking in multiple ways. By using the concept of documentation to make learning visible, the Reggio Emilia Approach does not attempt to fit the child to one particular assessment technique but fits assessment to the learner in multiple ways. Of the various types of early childhood assessment, most researchers suggest that only formative assessment can be identified in the Reggio Emilia Approach (Fyfe, 2012). However, the documentation processes in the Reggio Emilia Approach reveal students' skills and knowledge and can be used in many ways as an informal classroom assessment practice to gather "traces of learning (Gullo, 2004)." And although documentation, in the Reggio Emilia Approach, is meant to be an ongoing interpretation of student work, teachers practicing this approach have claimed that as an informal classroom assessment tool, this method of documentation is also helpful during times they need learning to be made visible (or summative) at the end of a unit or project (Fyfe, 2012). Formative or summative – all informal classroom assessment methods can be described as decision-making tools (Gullo, 2004).

Educators in Reggio Emilia have adopted what they refer to as, "the pedagogy of listening." They strive to respect young students' efforts to make meaning of their experiences (Rinaldi, 2006). In the Reggio Emilia Approach, assessment is focused on what the student can do rather than what the student cannot do. And educators engaged in this approach study and assess the ongoing development of children as well as the individual development of each child (Forman & Fyfe, 1998). This study will engage in the “pedagogy of listening” as an intervention by asking students to make sense of their experiences following a science inquiry. And through a story discourse that might employ multiple “languages,” asks students to make their learning visible. In the Reggio Emilia Approach, teachers listen with their third ear. This means that they practice listening for the implied meaning of student’s words (Forman & Fyfe, 1998). Using a “pedagogy of listening” and the theory of the “one hundred languages,” this research captured data from students’ stories of learning and compared the data to the teacher’s assessment to determine if the intervention provided an enhanced (or better) understanding of what the students learned.

Significance of the Study

The goal of this study is twofold. First, is to investigate existing teacher-driven summative science assessment practices. Second, is to present an alternative form of summative assessment to evaluate what students have learned. This study is important because flawed assessment practices could, theoretically, result in charting a course for learning that does not enable students to reach their full potential.

Enger and Yager (2009), asked two key questions that support this investigation. The questions include, “Should assessments tell us what students cannot do or what each student can do?” And, “How can assessments encourage and recognize inventive, imaginative responses that, although unexpected, are constructive and appropriate?” They go on to suggest that assessment practices should be linked to student outcomes and should mirror the ways in which students learn. Assessments should mimic students' learning opportunities, and ultimately, take students to higher cognitive levels. Traditional short answer assessments, whether they are true-false, fill-in-the-blank, multiple-choice, or other short answer-type tests are not student-centered enough for the inquiry-based, constructivist oriented, science lessons today.

Improving assessment tools, and teachers' ability to analyze student responses as windows into their thinking and understanding is at the heart of the challenge facing educational researchers today (Supovitz, 2012).

Research Questions

The focus of this study is as follows:

1. Do existing teacher-driven tests allow students to fully articulate what they have learned, following a science lesson?
2. Does the use of interview, as a form of summative assessment, allow students to fully articulate what they have learned, following a science lesson?
3. Can summative assessments be enhanced to provide students with more opportunities to demonstrate what they have learned?
4. How does the use of interview to assess student knowledge compare to existing teacher-driven summative classroom assessment methods?

Chapter 2

Literature Review

A Brief History of Assessment in Education

The evolution of assessment in early childhood education can be traced as far back as the late 1700s when Johann Pestalozzi, a pioneer in developing children's educational programs, wrote about the development of his son. (Irwin & Bushnell, 1980). Early publications by philosophers such as, Locke, Rousseau, and Froebel were extremely important to the evolution of assessment and the study of early childhood education in general. However, it wasn't until the late 19th and 20th centuries that the field began to receive the attention these early pioneers advocated for (Wortham, 2011). During this time scientists throughout the world used observation to measure human behaviors. Pavlov introduced his theory of conditioning. Binet developed the concept of a normal mental age and provided a plethora of research on memory, attention, and intelligence in children. Binet and Simon developed an intelligence scale that made it possible to differentiate the abilities of individual students (Weber, 1984).

It could be said that from the moment a child is born, assessment and evaluation play an important role in helping adults determine if there is a need for some type of academic, developmental, or behavioral intervention (Gullo, 1994). Within the first minute of life, children are observed, assessed, and given a numerical score based on the Apgar Scale (Apgar, 1953). The Binet Scale and testing of children in schools is linked to the French Minister of Education. In 1904, French educators indicated a need for a classification system for admitting, placing, and developing educational programming for children with special needs (Kelley & Surbeck, 1983). The Apgar Scale tested children's physical/health well-being and the Binet Scale marked the beginning of the educational testing movement (Gullo, 1994). A classroom test, or assessment, has maintained three key components over the past several decades – they are samples of learning, they consist of behavior, and they are given and taken under standard conditions (Trice, 2000).

By the early 1900s, the predominant view was that objective tests could be used by teachers to study and improve educational outcomes, as well as provide diagnosis and placement for individual students (Symonds, 1927; Thorndike, 1913). Additionally, measurement experts began to train teachers to make their own classroom tests. This resulted in a framework for testing – quizzes, formal testing, and grading, that we still use today. (Shepard, 2006).

In the late 1970s, the United States began mandating minimum competency testing at the state level as they took on a significant role in education reform policy. As the federal government took notice of what was happening in the states, interest in standardization in accountability and assessment across the U.S. resulted in the *No Child Left Behind Act of 2002*, signed by President George W. Bush (Ferrara & DeMauro, 2006). And although this type of accountability testing is managed externally by administrators and policy-makers, it was recommended that assessment information intended to guide the instruction of children should support teachers as they make ongoing and daily instructional decisions that guide student learning (Stiggins, 2002). The twenty-first century has provoked, and maintained, a perspective on assessment that is two-fold – public accountability and instructional design (Ferrara & DeMauro, 2006).

Assessment in Early Childhood Education

There are many reasons to use assessments in early childhood education. They are used in infant and toddler settings to predict, identify, or rule out developmental concerns. They are used in preschool settings to test for readiness. They are used at the community level to capture a broader picture of the skills and abilities of a particular student population. Diagnostic testing is used to identify very specific learning disabilities found in screenings. Generally speaking, assessments allow us to monitor student progress and guide instruction (NRC, 2008).

Throughout the educational continuum, including the early years, assessments are used to measure abilities, skills, and level of knowledge in a somewhat consistent fashion from student to student. Some assessments involve paper-and-pencil. Some involve a type of performance activity. Others involve the use of spoken language. Many types of assessments can be meaningful and allow educators to draw inferences about students' ideas, reasoning, and learning (McDevitt & Ormrod, 2010). When used properly, assessments can be valuable tools for students to demonstrate learning. These tools play a critical role in the educational process, from the classroom to policymakers. The new K-12 science framework clearly states that assessment will continue to be an essential component to science education as the field moves forward with integrated learning (NRC, 2014).

Classroom assessment may be formal or informal. Formal assessments consist primarily of high-stakes standardized tests that are held to specific administration and scoring procedures. These assessments may be criterion-based or norm-referenced. Informal assessments primarily include observations, documentation, and interviews that are less structured and not validated or tested for reliability. Informal assessments are often developed by teachers (National Collaborative on Workforce and Disabilities for Youth 2015).

Generally, informal classroom assessments fall into two categories – summative or formative. Summative is sometimes referred to as an assessment *of* learning, and formative as an assessment *for* learning (Looney, 2011). Additionally, a form of assessment that should not be overlooked in early childhood education is authentic (or alternative) assessment. This type of assessment uses natural settings and attempts to match the assessment model to the child's everyday activities (Bagnato, 2007).

Assessments provide a substantial portion of the data that teachers use to plan within the classroom. Assessment settings can range from informal questioning, to teacher developed tests, to annual high-stakes testing. Regardless of the context or timing of the test, assessments share the important role of informing teachers about the contextual understanding of their students and, subsequently, impact lesson planning and the overall effectiveness of pedagogy (Supovitz, 2012).

Selecting a data collection strategy is important. The results must accurately inform instruction, identify students who need additional support, and properly report progress to the students' families and other stakeholders. Assessments must be designed, implemented, and interpreted with respect to reliability, validity, and cultural bias (McDevitt & Ormrod, 2010).

Combining various ways to collect data, and using various ways to interpret and report data, creates many different methods and models of assessment. The most important thing for teachers to consider when designing, or choosing, an assessment is to consider the purpose of the assessment. (Harlen & Qualter, 2009).

Assessing younger students is much different than assessing older students, or adults. The primary reason for this is that young students learn differently. Young students are less likely to learn through abstract reasoning and paper-and-pencil activities. They tend to learn in experiential and hands-on ways such as touching, talking, listening, moving, and playing. Consequently, the expression of what they know may be best served in ways other than paper-and-pencil tests (Guddemi & Case, 2004). When evaluating young students, it is vital to know if an incorrect answer is because the student cannot read or if he/she interpreted the question wrong. Even correct answers should be scrutinized in this manner (King, 1975).

There are many specific assessment tools to choose from in early childhood education. The type of tool used depends on the intention of the evaluation. There are four collections of assessment tools that include – screening, informing instruction/monitoring progress, diagnostic, and program evaluation (Washington State Office of Superintendent of Public Instruction, 2008). Screening tools include those such as the Ages & Stages Questionnaires (Squires, Bricker, & Potter, 1997), the Battelle Developmental Inventory (Newborg, et al., 1988), and the Brigance Screens (Brigance, 1985). Tools that teachers choose to inform instruction might include the Devereux Early Childhood Assessment (LeBuffe & Naglieri, 2003), the High/Scope Preschool Observation Record (High/Scope, 1992), or the Creative Curriculum Developmental Continuums (Dodge, Colker, & Heroman, 2001). Diagnostic tools outnumber all other early childhood assessments and specialize in genres such as behavior, intelligence, language/literacy, motor skills, and special needs (Washington State, 2008). And lastly, widely used tools used to evaluate the classroom or program include the Early Childhood Environment Rating Scale – Revised (Harms & Clifford, 1982), the Classroom Assessment Scoring System (Rimm-Kaufman, La Paro, Downer, & Pianta, 2005).

Like standards, assessment will always be part of our current educational and political structure. Not only must teachers introduce and teach standards in a developmentally appropriate manner, they must also design and implement assessments in developmentally appropriate ways (Cress, 2004). Too often, assessments designed by teachers overemphasize memory for procedures and facts rather than aligning their assessment practices with their instructional goals related to depth-of-understanding (NRC, 2000).

The use of story as an assessment tool in the classroom can be a valuable alternative method of collecting data from students. It can be especially useful in research relating students who may not read well, resulting in socially inclusive research (Davis, 2007).

Summative assessment in early childhood education. Summative assessment is a way to measure what a student has learned at a particular point in time. Unlike formative assessment, where teaching and learning is happening during the assessment process, summative assessment is an attempt to take a snapshot of progress and provide adults with an interim report that organizes learners by a grade, or a pass/fail rating. Although a focus purely on summative assessment is an ill-conceived strategy when evaluating young learners, summative assessment is a useful tool for determining what has been learned and what progress needs to be made. Summative assessments are also useful for reporting progress to families and officials (Harrison & Howard, 2011). This type of assessment can be easily summarized in quantifiable ways and allow teachers to compare students over time and across settings (NRC, 2008). Many associate summative assessment with standardized testing exclusively, and fail to recognize the important contributions this type of assessment has on instructional planning, from the classroom to the district-wide level. Additionally, the data from summative assessments fill an important role in the grading process (Garrison & Ehringhaus, 2007).

It is often heard that formative and summative assessment practices cannot be used together, and this is a mistake (Theall & Franklin, 2010). Formative and summative assessment can, and should, be used in conjunction with one another rather than being viewed as competing forces. If formative assessment has been executed properly, then summative assessment data should indicate whether or not students have reached the curriculum goals (Ornstein & Hunkins, 2009). Oláh, Lawrence, and Riggen (2010) found that teachers relied on summative assessments to identify students' areas of weakness but did not use them to look deeper into students' conceptual understanding of the material.

Although some studies indicate that teachers favor formative assessment methods, there is no doubt that many continue to develop and administer summative tests in their classrooms. Unfortunately, when asked how many used the summative data to inform future instructional practices only 40 percent could confirm. And despite opportunities to use summative assessment data to inform pedagogy, it is clear that teachers are not taking full advantage of the potential that this type of information has on future planning (Hoover & Abrams, 2013).

Gardner, Harlen, Hayward, Stoart, and Montgomery (2010) contend that the popular distinction between formative and summative assessment – formative being what a teacher does during a lesson, and summative being what a teacher does at the end of a lesson or unit, is unhelpful. They argue that it is the purpose and use of the assessment, not the timing that determines whether it is formative or summative in nature. Because summative assessment practices can be used in formative ways, and formative assessment practices can be used in summative ways, it can become difficult for teachers to differentiate between the two methods. However, the flexibility of these methods is an asset to teachers. Teachers can use summative assessment prior to the end of a unit leaving them time to make beneficial modifications to instruction, and allowing students an opportunity to revise their thinking. Assessment only becomes formative in nature when the data is used to inform teaching or to impact learning. Teachers may also use carefully recorded formative data to report to parents in a more summative way (NRC, 2001).

There is evidence to suggest that we can do much more with summative data. Teachers put a lot of time and effort into grading homework, but they do not include the grades as part of the classroom and student dialogue that shapes learning. A numerical score or grade alone does not tell a student how to improve and, as a result, misdirected teachers lose an opportunity to enhance student learning (Black, Harrison, Lee, Marshall, & Wiliam, 2004).

Teacher-driven tests as summative assessment in early childhood education. For the past several decades, when teachers have conducted summative assessments, one of the most popular methods has been a paper-and-pencil test. This type of assessment is often faster and easier than other procedures teachers are trained to use. Although paper-and-pencil tests may be known to be best suited for short answer, multiple-choice, true-false, and matching tasks, they can also be used to assess critical thinking skills through essays, as well as asking students to synthesize their thoughts and ideas on paper (Ormrod, 2011). This type of testing gained momentum in the United States' educational system following the launch of Sputnik, when a plethora of sit quietly, paper-and-pencil activities were implemented so students could "catch up" with their Russian counterparts. The trend then continued in an effort to compete with students from Japan (Association for Supervision and Curriculum Development, 1993).

Paper-and-pencil tests are almost always used for summative purposes. One study indicated that when using paper-and-pencil tests, 91 percent of K-2nd grade teachers and 98 percent of 3rd and 4th grade teachers used them exclusively for summative assessment purposes. The researchers felt this was a significant finding because their teacher preparation programs had trained them to use all assessments as part of a formative assessment process, including the paper-and-pencil method. The teachers in the study contended that they needed to use paper-and-pencil summative assessments to get a clearer picture of which students were learning and which students were not (McNair, Bhargava, Adams, Edgerton, & Kypros, 2003).

An important quality of any assessment is that it is culturally competent (Meller & Ohr, 2008). Luykx et al. (2007) researched paper-and-pencil assessments in early childhood science and found that although the goal of avoiding cultural bias is one to continue to strive for. Test designers cannot be expected to produce culturally neutral paper-and-pencil assessments due to the inherent cultural nature of the tests to begin with.

Additionally, there is a great amount of cultural fluency that operates below the level of conscious awareness. Two important considerations related to paper-and-pencil assessments are linguistic influence and language development. Teachers are sometimes unable to see through linguistic interference to understand what students meant to write versus what they actually wrote. And limited proficiency in the language of assessment, or language development in general, can masquerade as limited content knowledge.

Even though most early childhood educators consider paper-and-pencil tests developmentally inappropriate when administered to young learners, this testing scenario often emerges when teachers think about measuring what has been learned. The risk to using this type of measurement is that students may not have yet developed the skills required to successfully record their responses, and therefore, may know the correct answer but not be able to articulate their understanding through a paper-and-pencil assessment activity (Kostelnik, Soderman, & Whiren, 2011). The Association for Supervision and Curriculum Development (Association for Supervision and Curriculum Development, 1993) also deemed paper-and-pencil assessment practices as developmentally inappropriate for young students, and suggested that this type of assessment tends to be inaccurate due to younger students' inability to understand or follow instructions well.

Authentic assessment in early childhood education. In addition to summative methods of evaluation in the early childhood classroom, there are a number of authentic (also referred to as alternative) approaches to assessing student progress (Gullo, 1994). Authentic assessment is defined as a developmentally appropriate alternative to conventional teacher-driven testing methods when evaluating students (Bagnato, 2007).

There are six dimensions in authentic assessment. The dimensions include – a commitment to gathering structured recordings/notes, making age or stage-related observations, providing ongoing monitoring, focusing on the student's natural competencies, ensuring that assessment is conducted by familiar people, and incorporating assessment into everyday routines (Bagnato, 2007). Contrary to authentic assessment, classroom tests are often conducted by unfamiliar "experts" relying on fixed test situations, using scripted materials and anticipating scripted child behavior (Bagnato, Neisworth, & Pretti-Frontczak, 2010). Authentic assessment investigates a student's natural behaviors through natural tactics by someone who the student knows. The most common method is by recording developmental observations (Bagnato & Ho, 2006). Trice argues that the authentic assessment process described in current literature may measure educational objectives in the most direct manner but using the term "authentic" to describe any type of assessment process makes an assumption that all other methods are unauthentic, and this is simply untrue (2000). Proponents of authentic assessment contend that the method *is* more valid than using rigid tasks and materials, or "tabletop" tests. And that in addition to observation, natural settings may also employ the use of rating scales, curriculum-based checklists, and interview inventories (2000).

Formative assessment in early childhood education. One broad definition of formative assessment is that it includes all activities that a teacher and/or student participate in to obtain data that can be used diagnostically to alter teaching and learning. This can include teacher observation, discussion, and analysis of student work through a variety of methods (Black & Wiliam, 1998). Definitions such as this are compatible with the view that formative assessment only transpires when the data (or evidence) is used to make needed changes in the pedagogical approach (Schneider & Randel, 2010). Upon closer examination, the exact meaning of formative assessment can become complicated. Some researchers consider all assessment in the classroom formative, and assessment outside the classroom summative (Antoniou & James, 2014). Some suggest that all classroom assessment *can be* formative but only if used for formative purposes, while others suggest that assessments completed for formative purposes can be used in summative ways. And some believe that formative assessment is exclusive to the use of an instrument or diagnostic tool (Pearson, 2005). And just as many researchers and educators agree that formative assessment is a process as opposed to a test that produces a score (Popham, 2008). What is clear is that no universally accepted definition of formative assessment exists, and there are multiple ways of conducting this type of assessment. What truly makes it formative is how the teacher uses the data to inform instruction (Popham, 2008). Formative assessment is largely an evaluation of an individual's teaching and a diagnostic assessment of teacher accountability (Trice, 2000).

Many aspects of early childhood education can be monitored by using formative assessment methods. In addition to monitoring student progress, formative assessment can provide feedback related to the equipment and materials used, curriculum activities, and teacher behavior (Wortham, 1990). The goal of a formative assessment strategy is to gather evidence concerning the effectiveness of the various components involved in the classroom curriculum as well as the teacher instruction. (Gullo, 1994). The process used for formative assessment requires careful planning, involving teachers and students. A test may be included in formative assessment but is only one part of a multistep process that may involve many different activities (Popham, 2008). In formative assessment, feedback is a vital component and is as one of the most influential factors related to student learning (Hattie, 2009). Throughout a unit, formative assessment is used to thoughtfully analyze progress and plan instruction. Ultimately, the unit ends and teachers switch from evaluating instruction to assessing individual students by using summative assessment methods (Trice, 2000).

Interviewing Young Students

With learned skills and experience, the dialogue and interaction between a researcher and a student participant can become a meaningful exchange that results in a productive partnership (Kortesluoma, Hentinen, & Nikkonen, 2003). Danby, Ewing, and Thorpe (2011) suggest that the first step to take when designing a plan to interview young students as part of a research project, is to include prior visits to their classroom. This allows the student to feel more comfortable and the researcher to enter the students' culture of communication (Christensen, 2004). To enter the students' culture of communication the researcher must create an acceptable form of dialogue and interaction that echoes the students' own communication and behaviors. A major task for the researcher can be finding a way to connect with the student participant, and establishing an environment where the child can begin to tell their story (Cameron, 2005).

Using story to create a narrative in early childhood education. Story has many uses in the education of young children. It benefits students with the acquisition of language and literacy (Kaderavek & Justice, 2002). Story provides students with a conceptual framework for thinking and allows them to shape their experiences into something they can understand (Collins, 1999). The use of story is pervasive in children's lives. Early on, infants are surrounded by stories that adults tell them, or they overhear being narrated to others. By the time children are two years old they are adding their own voices to the stories that have surrounded them (Engel, 1995). Throughout our lives, we depend on narratives. Sharing stories is a way to make sense of experiences that challenge the ordinary and require understanding (Emde, Wolf, & Oppenheim, 2003).

Over the last few decades education in general has taken an interest in the use of narrative. There has been a wave of narrative research and narratives are increasingly used as models for analysis as well as explanatory devices (Rutten & Soetaert, 2012). Jerome Bruner's theory of narrative as a specific 'mode of knowing' proved important for this growing body of work. Bruner's theory of narrative is an important part of his general theory of culture, mind and education. He argues that one of the principal ways in which our minds are shaped to daily life is through the stories that we tell and listen to (Bruner, 2006). And he is not the only scholar to emphasize that narratives convey a specific kind of knowledge. Wertsch (2002) also believes that narrative can be used as a tool in education. He describes narrative as being a unique instrument we have for representing settings, actors and events. Wertsch claims that a great deal of our thinking and speaking are fundamentally shaped by narratives. We are storytelling animals when it comes to recounting and interpreting our own and other's actions. Evidence of teaching and learning by using story can be found in many classroom environments. Young students with autism spectrum disorder learned appropriate social behavior through the use of social stories. The students who received the social story treatment were more successful than those who did not receive the treatment when it came to understanding the perspective of others and responding appropriately in social situations – even up to four weeks after the treatment was administered (Kuoch, 1998; Delano & Snell, 2006).

Curenton (2006) studied the use of story in preschool classrooms and found that, not only did it support language and literacy development, it contributed significantly to overall kindergarten readiness. The teaching and learning value of stories is endless. They can be used to teach children about difficult situations, moral lessons, and making good decisions. Story can even be used to teach children fluency in a targeted concept (Coleman & Weidner, 2009).

Booth and Barton (2000, p. 37) ask: "If story is a basic way of organizing experience, and if we search for our own stories in the stories of others, can narrative be a form of research that we can employ to examine education and our role in the teaching/learning process?"

The use of story to assess student knowledge in the early childhood classroom. Loris Malaguzzi, the founder of the Reggio Emilia Approach, believed that children have a hundred languages, or a hundred ways in which they can learn and communicate with their world (Edwards, Gandini, & Forman, 2012). Assessment should provide students with the opportunity to use all their languages (Glazzard, Chadwick, Webster, & Percival, 2010). The most important aspect of using story to assess student understanding is that stories have the ability to include the student's full voice in the assessment process (Lyle, 2000). And although young children have a complex vocabulary made up of thousands of words by the time they are five years old, researchers have also found that children understand much more than they are able to articulate (Snow, 2001).

One way that young students can tell the story of what they know is through drawing. Drawings are often utilized in early childhood education as a way for students to express their understanding and construct meaning. Drawings provide teachers with useful information and children's insights into the world around them. They also provide teachers with an opportunity to engage in dialogue that allows them to better understand students by listening to their stories. Drawings communicate the students' voices, which the teacher can then use in curriculum decisions. The teacher also consistently evaluates learning outcomes by listening to children, recording what they have to say, and reviewing the drawings for the purpose of modifying curriculum and instructional approaches (Chang, 2005). Drawings as stories are powerful when used as a tool to access the inner thoughts, feelings, and experiences of students and young students will often include a vocal narrative (MacDonald, 2009).

Interviewing, or asking students to orally deliver a story, is another method of using story to assess understanding. In the Reggio Emilia Approach, students are encouraged to talk about what they know, before and after a project. Eventually, these students learn to communicate at a metalinguistic level, as they not only talk about what they know but can explain how they represent what they know (Forman & Fyfe, 1998). The HighScope approach to early childhood education employs a "plan, do, review" system to teaching and learning. Students are asked to plan their own learning activity at the beginning of the day, engage in the activity, and then recall what they actually did during activity time. The "review," or recall, session provides students with an opportunity to think about what they did and what they learned as they share it with the teacher and the class (Wiltshire, 2013). Open-ended questions are best when interviewing young students, but they can also be challenging. The amount of information in young learners' memories may not be reflected within their responses during the interview. One way to address this challenge is by offering student participants the opportunity to use props to focus, organize, and expand their responses (Salmon, 2001).

Genishi and Dyson (2009) have conducted years of research relating to the importance of stories for young children. Through their research, they not only confirmed the value of listening to, and telling, stories but suggest that children can also benefit from "playing them out." Their research suggests that children naturally move from talking about an experience to acting it out, and that discourse is socially constructed and an instinctive part of the human experience. When students are given a variety of ways to dramatize ideas, they can negotiate through the ambiguity of meaning as well as experience better results when communicating their understanding (Varelas et al., 2010)

By using a narrative approach to understand what students have learned, teachers can protect the complexity of the learning process itself, by not allowing what is not tested to be lost. Quite simply put – much more student data can be collected (Carr 2001). When using storytelling as a tool to know what students know, the classroom not only benefits *from* the use of story, but *through* the use of story (Kallio, 2015).

A narrative approach to assessment will reflect what students have learned better than assessments with pre-set performance indicators. Although the direction might be difficult to predict, stories and narratives have the capability to capture the most elusive learning and development, that which is sometimes hidden from view (Carr, 2001). Short narratives provide a more accurate picture of student achievement than a numerical score or a letter grade. And learning outcomes deserve insightful answers, high levels of engagement, quality relationships, and opportunities to use the imagination (Eisner, 2005).

Talking to students about what they have learned is helpful in understanding their knowledge and it promotes a more meaningful relationship by allowing the teacher to get to know the student better (Glazzard, Chadwick, Webster, & Percival, 2010). Oral assessment data provides more detailed and reliable information than anything in a written context in the early years of development. Young learners' capacity to think and understand concepts far outweighs their ability to interpret what they've learned in written form (Fisher, 2004).

Though the use of story to assess children's knowledge in education is not prolific, there are many instances in early childhood education where researchers have come to know the value of using story to gain access into children's thinking and as a tool for teaching. Collins (1999) maintained that storytelling has many uses in primary classrooms. She concluded that stories provide a conceptual framework for thinking, which allows students to shape their experiences into a whole they can understand. Stories allow them to mentally map experiences and see pictures in their heads providing them with a model of language and thought that they can emulate. Carr (2011) found that a range of deliberate teacher strategies can provide opportunities to engage young students in revisiting and communicating what they've learned through the classroom practice of reflection.

Assessing Scientific Knowledge in Early Childhood Education

There are many things to teach and learn about science. There are more concepts than time to teach them. Aside from concepts, teaching and learning should also include the process of high quality inquiry-based science. The history of science should also be considered (Abell & Volkman, 2006). Even if all stakeholders agreed on *what* should be taught in early childhood science education, what is counted as evidence of learning would continue to be disputed (Abell & Volkman, 2006). Assessment in science education is concerned with knowledge of what aspects of science are important to measure, the methods by which concepts can be assessed, and the trade-offs involved in employing one method over another (Falk, 2011). Assessments in science education should emphasize the skills and scientific reasoning found in inquiry-based instruction and active learning (Anderson, 2011).

During the last two decades, researchers and leading science organizations such as the NRC have placed more emphasis on the importance of *how* children learn rather than the memorization of scientific concepts. The major goals of science education are to help children understand the modes of scientific inquiry and to foster their inquiry-based skills (National Science Education Standards, 1996). To assess student skills in science by today's standards, teachers must re-consider how they are going to collect the necessary data. Assessing all students at the same time is not recommended and focusing on a small number of children at a time should be considered. Additionally, assessing only two to three inquiry skills per day is recommended. Teachers are encouraged to use questioning strategies in science assessment, so they can clarify scientific actions with younger learners. By using questions related to their actions, teachers are able to better understand student behavior (Lee & Yoon, 2008).

The role of assessment should be to help students learn, not just measure outcomes. To use assessment as a tool to measure *and* enhance learning, the teacher must understand how to use the data. The quality, and use, of assessment data are both crucial. Sometimes, it is an easy task, and sometimes, as with activities such as inquiry-based science investigations, it is a complex task. Regardless of the level of complexity, opportunities for meaningful assessment occur daily within interactions, in conversations, through observations, and as part of traditional assessment. (Atkin & Coffey, 2003).

Science assessment methods, such as storytelling or drawing, have the capability to incorporate a variety of meaningful ways to collect data. In early childhood education, drawings play an important role in science inquiry. Not only do they represent ideas and help students with learning scientific concepts, they also help teachers understand what students know. When students are engaged in scientific inquiry their drawings can expose what they know, what they would like to know, and what they have learned (Chang, 2005).

The NRC (2000) recommends that that continually conducting new assessment research in science education will provide teachers with a variety of supports and learning opportunities that will, in turn, enable their classroom assessments to support learning. New assessment methods are needed to measure new kinds of science learning.

The use of story as a way to assess learning following science inquiry has not been widely researched.

The Reggio Emilia Approach

Over the past 50 years educators, parents, and citizens in the small city of Reggio Emilia, Italy have been building a public system for early childhood education. The public education system was founded by Loris Malaguzzi following World War II when the parents of young children in Reggio Emilia envisioned a brighter future for their city's youngest citizens. That vision included schools where each child's intellectual, emotional, social, and moral potentials would be carefully cultivated and guided in healthy, beautiful, loving environments (Edwards, Gandini, & Forman, 2012). The approach has become a widely recognized and revered method of early childhood education around the world (Edwards & Rinaldi, 2009; Mantovani, 2007; New, 2003). Though the Reggio Emilia Approach is most well-known as a preschool system, the city also operates a prolific infant/toddler system and classrooms ranging from pre-k to 5th grade at the Loris Malaguzzi Center (Biddle, 2015). Students who attend are from all socioeconomic and educational backgrounds and, by law, children with special "rights" are mainstreamed as well as given first priority (Edwards, Gandini, & Forman, 2012).

The Reggio Emilia Approach employs 12 principles (Rinaldi, 2013). They are:

- The image of the child – children are the protagonists of their own growth and have extraordinary potential for learning
- The hundred languages of children – children possess a hundred languages or a hundred ways of thinking, of learning, of understanding, of interacting and of doing
- Participation – the children, educators, and parents all contribute to the education process in ways that nurture culture, solidarity, responsibility, and inclusion
- Listening – active listening skills are developed and used by the children, parents, and educators and are used as the context for educational relationships
- Learning as a process of individual and group construction – every child is an active constructor of knowledge and learns strategies of research, discourse, and co-participation

- Educational research – research is shared between children and adults and is a priority from the classroom to the international level
- Educational documentation – documentation is integral to the Reggio Emilia philosophy and aids in making the learning processes visible and assessable
- Progettazione – progettazione is the process of planning and designing teaching and learning activities based on the trajectory of the interest of the children
- Organization – the organization of the work, the spaces, and the time that children and adults interact is a vital part of the values and choices of the approach
- Environment, spaces, and relations – the environment is considered the third teacher in the classroom and changes based on the projects and adults
- Professional development – ongoing learning is considered both the right and duty of each individual
- Assessment – a continuous process of finding meaning and value, and structuring the educational experience

A key contribution of the Reggio Emilia Approach to assessment in early childhood education is the students' use of graphic languages and other media to record, revisit, and represent their memories, ideas, hypotheses, observations, and feelings during the learning process. The Reggio Emilia experience has proven, beyond doubt, that very young students can explore and express their understanding of a phenomenon using a variety of ways, or "languages." These languages communicate their understanding and demonstrate their construction of knowledge more readily, more competently, and at a much younger age than we have previously recognized (Edwards, Gandini, & Forman, 1998). Graphic languages and other media emphasize student participation and culture and can be used as a way for young students to construct a framework for understanding and to tell their learning story to others (Carr, 2001).

In the Reggio Emilia Approach, the teacher balances active engagement with careful listening to each student's needs and interests (Edwards, 2002). The key assessment method in this approach is detailed child observation, known as documentation. The methods used to document learning are varied and the process includes contributions from both the students and the teacher. The student's work, photographs, material representations, and enactments are all used to demonstrate, or document, learning. Upon collecting this data, teachers may then engage in critical analysis to determine what has been learned (Lash, Monobe, Kursun Koptur, & Black, 2016). The concept of documentation as we know it today, in early childhood education, has been largely inspired by the schools in Reggio Emilia. Teachers know it as both a verb and a noun, an act and an artifact. Documentation that is inspired by the Reggio Emilia approach may be defined as a systematic act of collecting, interpreting, and reflecting on concrete traces of learning (Fiore & Rosenquest, 2010).

Chapter 3

Methodology

The purpose of this study was to thoroughly examine two methods of assessment following a science lesson in a second-grade classroom. And through the assessment methods, using qualitative analyses, determine the role that classroom assessment *type* plays in the process of providing students with the best opportunity to express what they know, or have learned.

The study examined two contexts in which informal, summative, assessments are expected to tell teachers what students know. More specifically, this study documented what second-grade students could demonstrate and articulate related to their understanding of a scientific concept. Students were asked to share what they learned by engaging in two different contexts of summative assessment strategies. The first context was the existing classroom teacher-driven assessment, and the second context employed an interview approach between the student and the researcher who asked the students to tell a "story" about what they learned. The ability of each assessment strategy to accommodate students' capability to share what they learned was studied and compared.

Theoretical Framework

The theoretical framework for this research was inspired by the Reggio Emilia approach, also known in the field of early childhood education as REA. The approach was developed in the municipal system of infant-toddler centers and preschools in Reggio Emilia, Italy. The success of the approach has brought it worldwide attention and is followed by practitioners from many countries (Edwards, Gandini, & Forman, 1998).

The pedagogy of listening. We can only help children find meaning in what they do, what they encounter, and what they experience by developing an interpretive theory, and a narrative, that gives meaning to the world around them. However, listening takes time. It is only when you really listen to students that you get into the time of dialogue and interior reflection (Rinaldi, 2012).

The relationship between documentation and assessment. Formal assessments are not part of the Reggio Emilia concept of documentation. However, documentation is a key principle of the Reggio Emilia approach and is used to discover "traces of learning." Some documentation practices used to trace learning include – observation, anecdotal records, and samples of children's work (Fyfe, 2012).

The hundred languages of children. Loris Malaguzzi, the founder of the Reggio Emilia approach, wrote that children have at least one hundred languages – at least one hundred ways of thinking, of doing, of expressing themselves. He said if we are going to document then we should not only just record but use the information we receive to make predictions. Malaguzzi believed that not only do children move around all the time physically, but that their minds, social exchanges, and languages are all in continuous motion as well.

A Qualitative Inquiry: Using an Interpretive Paradigm, Grounded Theory, and Narrative Methodology

An interpretivist paradigm is often used in qualitative research. Glesne (2011) suggests that the role of the social scientist is to access others' interpretations of some social phenomenon and to interpret the actions and intentions of themselves and others. She goes on to say that although an interpretive researcher does look for patterns in the analysis, they do not try to reduce multiple interpretations to number, or to a norm. Qualitative researchers observe, ask questions, and interact with their research participants. Schwandt (2007) suggests that unlike realists, who believe the world exists independently from the knower, qualitative researchers think like idealists – who believe that the world is always interpreted through the mind.

This study utilized a qualitative research approach. It is a narrative and interpretive type study that employed a methodological triangulation for analysis. Before discussing details about the site selection, participants, and procedures, I will provide an overview of social science research, and qualitative research specifically – what it means, and the relationship between qualitative design and this research. Of particular relevance to this study is the section on grounded theory and triangulation methods.

Paradigms of Social Science Research

To go into more detail as it relates to this project we must look at Glesne's (2011) paradigms of social science research – or ways of knowing. As seen in Glesne's *Becoming Qualitative Researchers* (2011), Usher (1996) defined paradigms as the following –

Paradigms are frameworks that function as maps or guides for scientific communities, determining important problems or issues for its members to address and defining acceptable theories or explanations, methods, and techniques to solve defined problems (p. 15).

During the Renaissance the power of religion, as related to ways of knowing, caved in to empiricism. Empiricists believed that explanations should be built on observations and experimentation rather than those read in religious texts. This period became what is referred to as the Age of Enlightenment and empiricism became the standard in research (Glesne, 2011). However, the first paradigm that Western researchers used as guide map was *logical positivism*. Logical positivism was a view that "knowledge was limited to what could be logically deduced from theory, operationally measured, and empirically replicated" (Patton, 2002, p. 92). Scholars believed that research could only be built upon itself, thus increasing the body of knowledge until something is known (Glesne, 2011). Thomas Kuhn challenged this way of thinking in 1962 when he published, *The Structure of Scientific Revolutions*. Kuhn, and other scholars, argued that "data and observations are theory-led, that theory is paradigm-led, and that paradigms are historically and culturally located" (Usher, 1996, p. 16).

Today, Glesne (2011) suggests that the theories and philosophies that guide the work of social scientists fall into four paradigms (also known as theoretical frameworks). They are – Positivism, Critical Theory, Poststructuralism, and Interpretivism.

Positivism. The term *positivism* refers to a theoretical framework that attempts to apply the natural science model (or scientific model) of research to investigations within the social world. Positivism assumes there are patterns, causes, and consequences in the social world just as there are in the natural science world (Denscombe, 2010). Positivism came from Auguste Comte, a nineteenth century French philosopher. Comte advocated for an approach to social science research that would be *positive* in its efforts to achieve reliable and concrete knowledge on behalf of the social world. Other terms used for this paradigm may include – *postpositivism,* referring to a less strict form of positivism, and *logical empiricism* (Glesne, 2011). Although positivism has been described and interpreted in many different ways, most scholars will agree on these key characteristics (Bryman, 1988):

- The methods of the natural sciences are appropriate for the study of social phenomenon
- Only those phenomena which are observable can be counted as knowledge
- Knowledge is developed inductively through the accumulation of verified facts
- Hypotheses are derived deductively from scientific theories to be tested empirically (the scientific method)
- Observations are the final arbiter in theoretical disputes
- Facts and values are distinct, thus making it possible to conduct objective enquiry

One would use a positivism paradigm if their "beliefs include a fixed reality external to people that can be measured and apprehended to some degree of accuracy" (Glesne, 2011, p. 7). The main purpose for using a positivistic paradigm would be to predict. Research associated with positivism includes experimental, quasi-experimental, and causal comparative (Glesne, 2011).

Critical theory. According to Jim Thomas (1993), a *critical theory* framework intends to take you beyond "what is" and toward "what could be." The critical theory paradigm tends to be political by nature and critical theorists speak on behalf of their subjects giving them a larger and more powerful voice summoning social consciousness and provoking societal change.

According the Glesne (2011), critical theory can also be referred to as feminist theory and critical race theory. Karl Marx, Max Weber, and Herbert Marcuse are just a few of the philosophers associated with critical theory. The central purpose within this paradigm is to emancipate. Associated research methodologies would include – critical ethnography, feminist research, participatory action research, and critical discourse analysis. Critical theory doesn't necessarily follow a specific set of methods but a few key characteristics of this type of researcher can be identified:

- See research as a political act because it not only relies on value systems but challenges them (Usher, 1996)
- Focus upon language or the tacit rules that regulate what can and cannot be said, who is blessed by authority to speak and who must listen instead, whose social constructions are valid and whose are not (Kincheloe & McLaren, 2003)
- Are often interested in the relationship between thought and action, and theory and practice and use processes that enable people to challenge the status quo (Higgs, 2001)

According to Kincheloe and McLaren (2003), a critical theorist is one who is concerned with issues primarily related to power and justice. Also, this type of researcher is concerned with the ways that the economy, and issues regarding race, class, gender, ideologies, discourses, education, religion, social institutions, and cultural dynamics interact with one another to build a social system. In this type of research, the methodology needs a back-and-forth interaction between the researcher and the participants. Typically, the researcher and participant/s share a common bond of sensitivity and respect for the interest and commitment that each bring to the project (Bradley-Levine & Carr, 2015).

Poststructuralism. Glesne (2011) labels poststructuralism in the same paradigm as postmodernism, postcolonialism, and post-Fordist. Although these frameworks can all be distinguished separately from one another, they also share the same underlying philosophies and are also used interchangeably. Poststructuralists consider human behavior and speech textual productions and work to deconstruct them in systematic ways that both include and exclude both ideas and people. Therefore, the central purpose of this type of researcher is deconstruction.

Poststructuralism can be a difficult paradigm to pin down. Simply put – poststructuralism researchers are concerned with uncovering the flaws, contradictions, and inconsistencies in everything we read (Shmoop Editorial Team, 2008).

Interpretivism. Of the four types of qualitative approach, this researcher will employ the interpretivist paradigm. This type of paradigm is what is known as *qualitative*. The researcher observes, interacts, and ask questions. In general, results are not reduced to numbers or to a norm. Other terms associated with interpretivism might include – constructivism, naturalism, phenomenological, and hermeneutical. Associated philosophers and theorists within this paradigm include – Geertz, Habermas, Husserl, Kant, Herbert, and Mead. Ideas related to the interpretivist paradigm reach as far back as Roman times but interpretivism, as a way to conduct research in the social sciences, should be credited to the work of Immanual Kant (Glesne, 2011).

The main role of the interpretivist researcher is to assess the actions and intentions of their participants, as well as study the interpretations of the participants. Many styles of interpretivism exist but all styles share common goals of understanding some human idea, action, or interaction – whether in a specific context or within the bigger picture (Glesne, 2011). According to Glesne, the interpretivist researcher's approach will include these key principles (p. 9):

- May result in hypothesis and theory
- Researcher as instrument
- Naturalistic
- Inductive
- Searches for patterns
- Seeks pluralism and complexity
- Makes minor use of numerical indices
- Descriptive write-up

According to Glesne (2011), the central purpose within an interpretive paradigm is, to understand. Fitting within that interpretive paradigm, the goals of this study include;

1. to compare two contexts related to summative assessment following a science investigation, and
2. to investigate the accuracy of the data collected from the assessments.

This research is highly interpretive, or *qualitative*. The researcher subscribed to the key principles of the interpretivist researcher listed above from Glesne (p.9).

In the following section, a detailed definition of qualitative research will be provided. This section will also explain the narrative approach that was used to collect data from the participants in this research.

Defining Qualitative Research

Defining qualitative research is difficult because as Denzin and Lincoln (2011) suggest it has different meanings for different people. They define qualitative research as being "multimethod in focus, involving an interpretive and naturalistic approach to its subject matter (p. 2)" and "involves the studied use and collection of a variety of empirical methods such as – case study, personal experience, introspective, life story, interview, observational, historical, interactional, and visual texts that describe routine and problematic moments and meanings in individual lives (p.2)."

Researchers who choose qualitative studies have a baffling number of choices when considering their approach. Tesch (1990) provided guidance on 28 approaches in four different categories. Crabtree and Miller (1992) discussed 18 approaches in which the researcher could choose to work within the domain of human life that was a primary concern (the individual, the social world, or the culture). Lancy (1993) organized qualitative research in several disciplines (anthropology, sociology, biology, cognitive psychology, and history). Wolcott (1992, 2008) offered a classification resembling a "tree" (p. 82) with multiple branches of strategy suggestions for collecting data. Denzin & Lincoln (2011) have organized and re-organized numerous types of qualitative research strategies over past years. According to Creswell (2013), the definition of qualitative research is:

Qualitative research begins with assumptions and the use of interpretive/theoretical frameworks that inform the study of research problems addressing the meaning individuals or groups ascribe to a social or human problem. To study this problem, qualitative researchers use an emerging qualitative approach to inquiry, the collection of data in a natural setting sensitive to the people and places under study, and data analysis that is both inductive and deductive and establishes patterns or themes. The final written report or presentation includes the voices of participants, the reflexivity of the researcher, a complex description and interpretation of the problem, and its contribution to the literature or a call for change (p.44).

Qualitative research stresses implications that are not measured by quantity, amount, intensity, or frequency (Denzin & Lincoln, 2011). Marshall and Rossman (2014) described three distinct purposes for qualitative research. They are – 1) to *explore*, 2) to *explain*, and 3) to *describe*. They go on to say that alternative descriptions to those purposes could potentially be – 1) to understand, 2) to develop, and 3) to discover. To *explore*, according to Marshall and Rossman (2014), means to identify themes, patterns, and categories of the participants, as well as their relationships. To *explain* attempts to identify the causes of the patterns/themes and their relationships. To *describe* is the documentation of the phenomena of interest, and in the case of this study the teacher-driven assessments and the student narrative stories of learning collected by the researcher.

The narrative feature of qualitative research. Creswell (2013) groups qualitative research into five categories. The categories are – 1) narrative, 2) phenomenology, 3) grounded theory, 4) ethnography, and 5) case study. Though this research could potentially be regarded as narrative *or* phenomenological, narrative research is a closer match given that the main focus is primarily on the individual learning story of each participant. While phenomenology also focuses on the experience of the participant, it is a better fit for studies interested in the common, or shared, experiences of many participants (Creswell, 2013). "A narrative researcher can take a literary orientation for his/her analysis…" (p. 189). Similar to a method this research aims to use, Creswell and Ollerenshaw (Ollerenshaw & Creswell, 2002) used story, in a science education project, told by fourth graders.

This research is starkly narrative by nature because the researcher collected narrative information through an interview technique that contributed to each participant's *story of learning.* This included comparing how each participant performed in each of two methods of assessment following a science investigation. Creswell (2013) suggests that numerous types of narrative approaches have emerged over the years. The type of narrative that this research employed is an *oral history*. An oral history consists of collecting the reflections of events from one or several individuals (Plummer, 1983). Narratives often have a specific contextual focus such as stories told by teachers or students in the classroom (Ollerenshaw & Creswell, 2002).

According to Denscombe (2010), narrative in research must hold the following qualities:

- Have some specific purpose, which could be a moral message or an account of a personal circumstance
- Contain a plot line linking the past to the present, where there is some sense of development and/or change over time
- Involve people, a human element referring to feelings and experiences within the context of social events and human interaction

Narrative analysis is "concerned with the way that stories are used by people as ways of representing themselves and their lives to others" (p. 291).

Jerome Bruner wrote that if we do not tell about our experiences, then we do not exist (Bruner, 1991, 2004). Communication in the classroom is essential. It is through documentation that the processes of the student and teacher's research and thinking can be seen (Edwards, Gandini, & Forman, 2012).

When narratives are collected by interview, the protocol is to ask broad, open-ended, questions such as "Tell me about…" and allowing the participant to tell his or her story (Glesne, 2011, p. 185). Student participants in this dissertation research were asked a broad, open-ended question – "Can you tell me a story about what you learned in science today?"

Grounded Theory

Grounded theory was used to support this research. The intent of grounded theory is to develop or discover a theory that may help explain a particular practice or provide a platform for additional research (Corbin & Strauss, 2007). Ethnographic research shares some basic principles with grounded theory but examines entire culture-shaping groups at once (Harris, 1968). This research project does not meet the threshold to be considered ethnographic in nature. It could be argued that this research is a case study – an exploration of a real life, contemporary bounded system using interviews, audiovisual material, documents, and reports (Yin, 2009; Creswell, 2013). However, the focus of this research isn't as concerned with such an in-depth understanding of a single case as it is with sampling several learning stories that could impact future research for a community of learners. Additionally, because this study is not interested in a longitudinal data collection, and will not be an intense study of the classroom or the student participants, it does not qualify as a case study. Because the idea of using story as a summative evaluation tool is a unique concept, a grounded theory approach allows the researcher to build the research as it plays out rather than having to fully design the plan prior to data collection (Charmaz, 2001).

J. S. Lowenberg's classification (1993) of grounded theory best supports its use in this research – he states that grounded theory is a type of interpretive research best described as a form of symbolic interaction.

Grounded theory was first introduced by Glaser and Strauss in their 1967 book, *The Discovery of Grounded Theory.* Their work provided researchers with strong rationale for using qualitative research to develop theories (Goulding, 1998). One of the main principles is that the theory evolves during the research process itself. The relationship between the data collection and the data analysis is constant and continually developing (Glaser, 1992; Glaser and Strauss, 1967; Strauss, 1991; Strauss and Corbin, 1990, 1994). Grounded theory is a methodology that is used to construct theory where not much may be known, as well as provide new insight to existing knowledge. Grounded theory is an interpretivist paradigm where words, gestures, expressions, and behavior are all valuable and considered to be an important part of the overall experience between researcher and participant (Goulding, 1998).

Grounded theory is a popular methodology for those conducting research in small-scale projects using qualitative data to study human interaction, or by those conducting exploratory research in particular settings (Denscombe, 2010). Denscombe defines the grounded theory approach as a method used to generate theories, not test them. He goes on to say that this method should emphasize the importance of empirical field work and link revelations to the real world. Grounded theory is most useful to four types of research – 1) qualitative, 2) exploratory, 3) human interactions, and 4) small-scale research. This dissertation fits well within all four of Denscombe's categories of research due to its highly qualitative and interpretive design, and that the site selection involves human interaction that will be limited to two second-grade classrooms. Denscombe also believes that grounded theory is beneficial to a new researcher due to the systematic way of analyzing qualitative data using a constant comparative method.

The constant comparative method. The grounded theory approach employs a constant comparative method to analyze the data (Glaser & Strauss, 1967; Corbin & Strauss, 1990; Strauss & Corbin, 1994). Using this method, the researcher is constantly comparing and contrasting new codes and concepts to existing versions as they emerge. This type of approach allows the researcher to stay close to the data as well as what is actually happening in the field (Denscombe, 2010). Analysis should always include:

- Coding and categorizing the raw data (e.g., interview tapes)
- Constantly comparing the emerging codes and categories with the data
- Checking them out against new data specifically collected for purpose
- Generating concepts and theories that are thoroughly grounded in the data (p. 116)

The constant comparative method is a credible and trustworthy approach to data analysis that ensures the results, and any theory generated, remains true to the data that was collected. Additionally, any theory that is developed "remains grounded in empirical reality" (Denscombe, 2010, p. 116). Creswell (2013) offers several accuracy, or validation, strategies. Four of those strategies fit nicely within the framework of this research. They are – prolonged engagement and persistent observation in the field, peer review or debriefing, rich and thick descriptions of the data and themes, and external audits that allow outsiders to examine the product and process being generated within the research (pp. 251-253).

Constant comparative method uses multiple back and forth movements sometimes referred to as iterations, and the process continues until the researcher reaches saturation. Saturation is the point in which nothing new is being learned or observed from the data (Robert Wood Johnson Foundation, 2016).

In this section, qualitative research was defined with an emphasis on grounded theory and the constant comparative method. These approaches are vital to a high quality, qualitative, research project and are the appropriate methods for this study. In the following section, credibility and trustworthiness will be discussed in qualitative research design. Additionally, the importance of using the triangulation method to maintain the honesty and integrity of the data collected and analyzed will be discussed in detail.

Credibility and Trustworthiness in Qualitative Research

Due to the nature of qualitative research, it is not possible to verify it in the same way as quantitative research. However, every attempt should be made to demonstrate the credibility of the work (Denscombe, 2010, p. 298). According to Denscombe (2010), there are four viewpoints that are typically used to judge credibility. They are:

- Validity – referring to the accuracy and precision of the data.

- Reliability – referring to whether an instrument is neutral in its effect and consistent across multiples instances of use.
- Generalizability – referring to the idea of applying the results/findings to other examples of the phenomenon.
- Objectivity – determines that the research is impartial, neutral, and free from the researcher's influence on the outcome.

Denscombe goes on to say that the credibility of qualitative research is not easily judged using the abovementioned criteria. First, and foremost, it is impossible to replicate a social setting. And, secondly, the researcher is so intimately involved with the collection and analysis of the data that it would be extremely difficult for another researcher to produce an identical method and conclusion.

The standard criteria used to judge credibility in quantitative, or mixed methods, research does not fit within a qualitative study, so we must look to Lincoln and Guba (1985) for criteria relating to credibility and trustworthiness that are comparative with validity, reliability, generalizability, and objectivity in quantitative research.

Establishing trust. Trustworthiness is similar to the idea of validity and reliability in a positivist study. According to Lincoln and Guba (1985) there are four main criteria for establishing trust in a qualitative study. Those criteria are – *credibility, transferability*, *dependability, and confirmability.* I will address credibility last so that triangulation can be addressed in detail.

Transferability, is the qualitative researcher's imaginative process of using information collected to ask, to what extent the findings *could* be transferred to other situations? This is a comparative response to generalizability in quantitative research where the question asked is, to what extent are the results likely to exist in other situations (Denscombe, 2010)?

Dependability, as described by Lincoln & Guba (1985) is a way to ensure that research reflects explicit procedures and decisions that other researchers can see and evaluate (Denscombe, 2010). Denscombe states that it is only through this detailed information that it could be determined that there is a possibility of another researcher coming up with similar results. Lincoln and Guba (1985) refer to this account of very detailed records as an audit trail.

Confirmability attempts to capture the quantitative concept of objectivity (Marshall & Rossman, 1999) and if the results of one study can be confirmed by another (Lincoln & Guba, 1985). According to Schwandt (1997), confirmability should be able to establish that the data and interpretations of a research inquiry are not imagined. The researcher should have data that addresses the question/s and phenomenon/s that the researcher has described in the study. Lincoln and Guba (1985) propose that audit trails also be used to achieve confirmability.

According to Lincoln and Guba (1985), there are five distinctive ways to increase the *credibility* of a qualitative research study (Lincoln and Guba merged prolonged engagement, persistent observation, and triangulation into one category. For the sake of clarity, they will be separated into three distinct categories bringing the total to seven in the following description):

1. Prolonged Engagement – the investment of time required to learn about the culture and participants of the study to minimize misrepresentations (being in a new place can draw unwanted attention to the researcher and

prolonged engagement minimizes this risk resulting in authentic findings), and build trust (Lincoln & Guba, 1985).

2. Persistent Observation – this tactic permits the researcher to identify the elements that are important to the study, and then to further investigate them through additional observation (Lincoln & Guba, 1985). Additionally, this allows the researcher to discard irrelevant elements (Eisner, 1975).
3. Triangulation – this process improves the probability that the results reported, and the analyses made are trustworthy. This is completed through the use of multiple sources, methods, investigators, and sometimes theories (Denzin, 1978).
4. Peer Debriefing – in peer debriefing the researcher will share data with a disconnected peer so that aspects that the researcher may have missed can be discovered. These sessions provide the researcher with an opportunity to clear the mind of emotions and feelings that may be clouding their judgment. Lincoln and Guba (1985, p. 308) refer to this as, “catharsis.”
5. Negative Case Analysis – the researcher uses hindsight to revisit, revise, and refine the hypothesis. A negative case analysis will eliminate any negative outliers/exceptions through the constant revision of the hypothesis and research questions until a perfect fit occurs (Lincoln & Guba, 1985).
6. Referential Adequacy – this is established when the researcher submits a portion of the data to an archive, so it can be examined at a later date and

compared to the critiques derived from the collected data (Lincoln & Guba, 1985). The archived data serves as a reference point for future data analysis. The interpretations can be tested for accuracy (Lincoln & Guba, 1985).

7. Member Checks – member checks occur when the researcher returns to the participants and asks for their feedback as it relates to the accuracy of the results/findings (Schwandt, 1997; Lincoln & Guba, 1985).

The two primary approaches used to infuse this qualitative research with credibility were – *peer debriefing*, and a significant usage of triangulation.

Peer debriefing in qualitative research. Onwuegbuzie, Leech, and Collins (2008) suggest that six types of debriefing techniques exist in qualitative research. They define debriefing the researcher as – an individual not involved in the study interviewing the researcher to provide him/her with the opportunity to reflect on the presence, or absence, of bias within the study. See Appendix A for the peer debriefing report related to this study.

The six ways that credibility is ensured through debriefing techniques include – debriefing the participants at the end of the study, debriefing the gatekeeper (the individual who is giving the researcher access to the site and/or participants), debriefing via multiple researchers involved in the same study, debriefing the focus groups participating in the study, *peer debriefing*, and *debriefing the interpretive researcher* (Onwuegbuzie, Leech, & Collins 2008). This study employed the peer debriefing technique and due to the nature of the study (highly interpretive) the sixth technique, debriefing the interpretive researcher, was automatically utilized. Lincoln and Guba

(1985) suggest that peer debriefing keeps the researcher honest by asking difficult questions related to the procedures and analyses. The "peer" can be any person not directly involved in the research – in other words, a "disinterested peer" (p. 308). The peer will serve as a second set of eyes to review and explore the data, as well as ask the researcher questions related to the procedures and practices used. Lincoln and Guba go on to say that when executed properly, peer debriefing is unquestionable.

The final type of debriefing discussed here, and that which was used in this study, is – debriefing the interpretive researcher. This type of debriefing is similar to peer debriefing but focuses solely on a formal interviewing method with questions planned in advance by the peer interviewer (Onwuegbuzie, Leech, & Collins, 2008). The questions are concerned with bias and authenticity within the study and are determined by the peer interviewer once he/she becomes familiar with the research project. The peer debriefing report can be seen in Appendix A.

Triangulation in qualitative research. According to Denzin (2006), there are four types of triangulation. Data triangulation consists of examining the consistency of different data sources from within the same method, such as changing the points in time the information is collected. Investigator triangulation is the act of using multiple analysts to review the data, or multiple observers to record the data, or both. Theory triangulation uses multiple theoretical perspectives to examine and interpret the data. However, Denzin's methodological triangulation is the best match for this study. Methodological triangulation involves the process of investigating the consistency of the results generated by multiple and various data collection methods (Denzin, 2006). Rather than seeing triangulation as a method for validation or verification, qualitative researchers generally use triangulation to ensure that their analysis is vigorous, comprehensive, and profound (Robert Wood Johnson Foundation, 2015).

Denscombe (2010) dissects methodological triangulation into two categories – between-methods and within-methods. Within-methods triangulation operates on the idea that comparisons using similar methods can prove the accuracy of the results. The school of thought here is that if similar methods produce the same findings then they must be considered accurate. The method best suited for this dissertation was the between-methods form. This method is the most common for social researchers. In between-methods triangulation the researcher uses multiple methods of data collection so the results from one method can be contrasted to another. Denscombe asserts that the use of differing methods allows a comparison of findings that produce these two main benefits:

- Findings can be corroborated or questioned by comparing the data produced by different methods.
- Findings can be complemented by adding something new and different from one method to what is known about the topic using another method (pp. 346-347).

This research is concerned with providing students with an opportunity to successfully demonstrate what they have learned, or know, about a scientific phenomenon. The method chosen for this study used more than one lens to see inside learning, or to gain an understanding about what students have learned. Triangulation is a practice that allows researchers to view things from multiple perspectives and supported this research by painting a fuller picture of data related to assessment in elementary science classrooms. The elements of triangulation for this research included interviews (student and teacher), teacher-driven assessment methods, and researcher observations.

Glesne (2011) suggests that proving the trustworthiness of your data includes being up front about the limitations of your study. The researcher's responsibility is to do the best that he/she can under the circumstances that are created during data collection. In other words, being open about all aspects of your research in terms of what went right and what went wrong will help the reader connect to your work. She also suggests that it is important to organize and document all data – including any thoughts you might have while collecting the data. Glesne goes on to suggest keeping all versions of the codes generated, and organizing them in chronological order. Lastly, keep all versions of your work and show your working documents.

In the preceding pages, I have attempted to define qualitative research and how credibility and trustworthiness can be achieved. The goal in providing such a detailed description is so that the reader would have a better understanding of how qualitative research is conducted. In the next section, I have explained the methods and procedures I used in this study, beginning with site selection and setting.

Site Selection and Setting

This research was conducted in two second-grade classrooms in an urban public-school district in the Midwestern United States. Over 21,000 students are enrolled in the district, of which 75% participate in the free and reduced lunch program. The racial make-up of the student population is – 41% Black, 38% White (not Hispanic or Latino), 11% Hispanic or Latino, 0.5 % Asian/Pacific Islander, 0.09% American Indian and Alaska Native (alone), and 8.7% other (including multi-racial). The largest minority group is Black, and the total minority population is approximately 62%. The racial make-up of the site selection city is different than the student population at – 58% White (not Hispanic or Latino), 27% Black, 8% Hispanic or Latino, 3% Multi-racial, 1% Asian, 2% Other, 1% American Indian, and .07% Native Hawaiian/Pacific Islander (City-Data, 2015).

The school and classroom were chosen *purposefully.* Purposeful selection allows for an in-depth study of information-rich situations in which one can learn a great deal about matters related to the research (Patton, 2002). Random sampling does not work in the context of this research since it is qualitative by nature and the researcher is not working with a population large enough to obtain a statistically representative sample that may result in widespread generalizations (Patton, 2002). Empirical generalizations of this type are typically applied to settings and populations beyond those specific to the study (Ritchie & Lewis, 2003). It is not a goal of this research to apply empirical generalizations, but to draw conclusions from single study data which may then be used in developing a wider theory in the future. This is a different type of generalization known as, theory building (Ritchie & Lewis, 2003).

Due to the nature of the topic, the research for this project did not call for a specific type of classroom site. Although demographic data has been collected, and an attempt to find the most diverse population was made, the grade level of the student participants fulfills the goal for site selection based on the purpose of the study. However, it is important to emphasize that the diverse ethnic and socio-economic population of the city and school district was an appealing and strong factor in the site selection. The researcher recognizes that data related to specific demographics could play an important role in additional research on this topic in the future. The school will be referred to as the Midwestern Public School, or MPS, within this research.

Participants in the Study

Two second-grade classrooms were purposefully chosen for data collection. The teachers and the science support specialist were awarded a stipend of $200 upon successfully completing their roles in the research study. The parents of the participating students were awarded a $20 gift card to Wal-Mart upon successful completion of their students' roles. The gift card was mailed directly to them by U.S. Mail to ensure their privacy and ensure they received the card. Additionally, all students in the classroom (participants and non-participants) were given small gifts after each day they were interviewed. Following the first interview, all students received a jigsaw puzzle of the United States of America. Following the second day of interviews, all students chose from a variety of National Geographic science books. All gifts were purchased at the local Dollar Tree store for no more than 1 dollar each.

Incentive payments are a worthwhile consideration in qualitative research given the demanding and intrusive nature of the data collection (Wilmot, 2005). Incentives can be used for participant recruitment purposes in many situations without any ethical concerns as long as all other ethical criteria are met. If the research meets all other human subjects' ethical criteria, then it should be considered benign. Incentives only become problematic when coupled with the following factors:

- Where the subject is in a dependency relationship with the researcher.
- Where the risks are particularly high.
- Where the research is degrading.

- Where the participant will only consent if the incentive is relatively large because the participant's aversion to the study is strong.
- Where the aversion is a principled one (Grant & Sugarman, 2004).

All ethical considerations related to incentives have been thoroughly examined and it has been concluded by the researcher that no ethical burden exists based on the design of, and protocols being used in, the data collection.

The two classrooms were purposefully chosen by a referral from a former colleague with an elementary science and research background. It is considered quite reasonable for a researcher to choose a convenient site to collect data when there are multiple equally valid possibilities (Denscombe, 2014). Once permission to reach out to school district staff was granted through a letter from the district superintendent, the former colleague referred several science support specialists and second-grade teachers to the project. The potential participants were given the researcher's email address. The order in which the science support specialists and teachers reached out to the project was the order in which they were scheduled to be interviewed for participation. The potential participants were contacted through email, and provided a brief overview of the project. The first science support specialist and first two teachers to express interest were from the same building which became an appealing and convenient option for the researcher. Once their interest was confirmed through email correspondence, an onsite meeting was scheduled. The meeting took place at the teacher participants' school in the science support specialist's classroom. The meeting lasted about one hour. Participation packets were given to all three participants that included an education background questionnaire and an informed consent form. The first science support specialist, and the first two second-grade teachers interviewed, and met with, were chosen to participate in the study. However, a wait list was also established in case someone would need to drop out of the study.

The participants in the study included one principle investigator and one graduate student investigator (referred to as researcher in this study) from a Midwestern higher education institution, one science support specialist, two second-grade teachers in their natural classroom environments, and 26 second-grade students in their natural classroom and library environments. The average age of the student participant was 8 years old. The classrooms are mainstream environments, meaning that children with special needs and/or those who are gifted may also participate in the study. However, information related to students who are gifted, have special needs, and/or disabilities was not collected as part of this study.

Detailed demographic data on student participants was collected using a parent/guardian questionnaire filled out at home (see Appendix B), and a Student Ranking Achievement Form filled out by each teacher (see Appendix C). Teacher demographic data, experience, and educational background were all collected during the initial interview/meeting (See Appendix D) prior to being confirmed as teacher participants. Both sets of data will be discussed in more detail in Chapter 4. The overall demographic data of the student population differed from that of the location city. Minorities enrolled in the school system represent approximately 62 percent of the student body. Well over half of the students in this school system are considered economically disadvantaged and over 75 percent participate in the free and/or reduced lunch program.

To summarize, participants in the study included one science support specialist, two second-grade teachers, and 26 students from one elementary school. The timeframe for the researcher's intervention was set to four sessions, consisting of two sessions on two different days, one week apart. Pseudonyms and/or codes were assigned to represent all participants and places described in this study.

Procedure/Methods

Consent was obtained from Human Subjects to conduct the research study. Following Human Subjects approval, contact was made with the colleague who would refer potential teacher participants to the study. Once the teachers were chosen, contact was made with their building principal to seek further approval to conduct the research in their building. The researcher also presented the building principal with the superintendent's letter giving permission to recruit teachers within the district. Once the building principal was briefed and permission was granted, the meeting was set up with one science support specialist and two second-grade teachers. During the meeting, the research project was thoroughly explained. It was explained to the science support specialist and the participating teachers that their key roles were to teach the science lesson in their natural environments, and administer their own teacher-driven classroom assessments following the lesson. Although the participating teachers would not behave differently, or use different lessons, methods, or assessments, there are three features of the research process that may be different than their normal routines: 1) they may be asked to implement the science lesson on a different day or at a different time, 2) the student participants will be divided into two groups for assessment, and 3) the participating teacher will need to administer the teacher-driven assessment two times for every lesson, once to each half of the student participants. Additionally, the environment would be altered by the presence of an audio recorder during the sessions.

The researcher visited the school prior to the implementation of the research study for two reasons – to scope out the best area to conduct the interviews and test the equipment, and to familiarize the student participants with the presence of the researcher. No recordings were initiated during the practice and familiarization session. During the practice session, the researcher spoke to the student participants about the audio equipment and explained that their work is being recorded so that researchers and teachers can better understand how students share what they have learned.

Prep work at the site. A letter of permission to conduct the study was obtained from the superintendent of the MPS. A consent agreement was signed by the teachers and science support specialist. The appropriate permission and consent forms were distributed to parents on behalf of the second-grade students who agree to be interviewed, audiotaped, and participate in the study. Only those children whose parents/guardians returned a consent form could participate in this study. Children who did not participate in the study did not leave the classroom to be interviewed. The non-participating student assessment scores, from the routine teacher-driven assessments, were not requested or recorded in any way as part of this study.

Instruments/measures. A total of five instruments were used in this study. Based on the qualitative nature of the study, the *researcher functioned as the first, and primary, instrument*. It is through the researcher's own experiences and methods (observations, interviews, recordings, etc.) that he/she learns and understands about the investigation and/or phenomenon in question (Lincoln & Guba, 1985; Marshall & Rossman, 2014).

Lincoln and Guba (1985) provide several reasons as to why the researcher is an important instrument in a qualitative study:

- The researcher can respond to and sense all the personal and environmental cues that exist allowing him/her to interact directly with the situation.
- The researcher is able to generate questions and hypotheses as data is being collected, and immediately test those questions and hypotheses with participants.
- The researcher can quickly summarize the data being collected and request feedback or clarification from the participants.
- The researcher can pay close attention, study, and include potentially important data from outliers rather than discarding them as happens in quantitative research.

The *second and third instruments used in the study were the existing assessments (for two lessons) developed by the science support specialist and used by the two participating classroom teachers* (see Appendices E and F). These methods were analyzed in detail once the data collection was complete. It was the goal of this researcher to use the existing teacher-driven method/s of assessment so that a comparison could be made between the use of interview and what was currently being implemented to understand what students have learned following a science investigation. If the existing method would not have been a match for this research (i.e., the teacher-driven assessment is too closely related to the researcher's method), the researcher was prepared to find alternative participants. These assessments were administered to one half of the participating students immediately following the conclusion of the first science investigation in each of two classrooms. The second half of the student participant group were given the teacher-driven assessment after they completed a different assessment task – described below as the fourth instrument. The order that the student participants in each classroom took the teacher-driven assessments was reversed following the second investigation scheduled on a different day (one week apart).

The *fourth instrument used in the study was the researcher's interview*. Each student was interviewed and asked to tell the researcher a story based on what they learned about the scientific concept being explored in the day's lesson. This practice was referred to as a "story" interview. Telling stories is essentially a meaning-making process and at the heart of what it means to be human is the ability of people to symbolize their experiences through the use of language (Seidman, 2006). Elliot Eisner (2005) said that teachers are in a position to interpret the quality of student questions and the insight of his/her answers. Narrative inquiry should be a source of data for understanding what young students are learning, and it is up to the researchers to design practices that involve short narratives that could be used to provide a much more complete picture of achievement.

During the interview process, some students chose to elaborate (or clarify) their learning story by using props, or materials, to assist them with articulating or demonstrating their thinking. A detailed list and description of what props were used by the student participants are included in the results (see Appendix E). A protocol for the story/post-lesson interview was developed (see Appendix F) and used for a total of four individual student interview sessions (two for each student), across two classrooms. This instrument was used as a guide during all four interview sessions conducted by the researcher.

And lastly, the *fifth instrument consisted of the audiotaped sessions* of the student story interviews. These sessions were reviewed and coded by the researcher in the days following the implementation of the classroom activity.

Data collection. The data collection included obtaining the teacher-driven science assessment results from the classroom teacher, as well as the interpretation of the recorded summative (story interview) sessions conducted by the researcher. The researcher also used anecdotal records transcribed during the story interview process. Anecdotal records are important because they have the capability to capture non-verbal data that may be important to the eventual write-up of the results. These observations might include things such as changes in the environment and/or the behavior of the participant (Glesne, 2011).

More specifically, the participating students were interviewed, observed, and audiotaped demonstrating what they learned following a science investigation. The participating students were also administered a teacher-driven assessment implemented by their regular classroom teacher. This type of data collection required qualitative analytic coding, that is, through observation – discerning themes, patterns, processes, making comparisons, and building theoretical explanations (Glesne, 2011).

Once the researcher identified the participants and prepared the instruments, the data collection in the MPS classroom began. The researcher contacted the science support specialist and the two classroom teachers to set dates for the data collection events. Prior to these dates, a meeting with the science support specialist, and teachers, was offered in the event they had any questions or concerns before beginning the project. It was unanimous, that an additional meeting was not necessary.

Prior to beginning the on-site data collection activities, the researcher collected all consent forms from the participating teachers, parents, and students. The forms and letters explaining the research to the parents were distributed following the initial meeting with the confirmed participating science support specialist, and teachers.

A researcher protects the anonymity of the participants by assigning numbers or aliases to all individuals involved (Creswell, 2013). All participating students in each class were assigned an anonymous identification code starting with SP = Student Participant (1-14), and C = Classroom (1 or 2). For example – the student in Classroom 1 who was interviewed first was Student Participant 1 (SP1) – meaning their code would be – SP1C1. Additionally, the participating students were given large round green stickers to wear on the front of their shirts and the non-participating students were given large round red stickers to wear on the front of their shirts. This was to help the researcher protect the non-participating students from inadvertently participating in the study, and used as a safeguard to avoid collecting data on non-participating students. When distributing the stickers to the students the researcher truthfully explained why everyone had a sticker, and why the stickers were different colors. The researcher also explained that everyone would receive a gift of appreciation whether or not they were allowed to participate in the interview. Non-participating students were not interviewed by the researcher and no data were acquired related to the non-participating students.

The teachers reminded the students about the research project prior to the arrival of the researcher, and prior to the beginning of their science investigation. Since the science lesson was part of their routine schedule, and the lesson itself was not being analyzed, the researcher opted not to be present in the science classroom during the investigation. Furthermore, the researcher felt that it was important to know as little as possible about the lesson so not to lead the student participant responses. The researcher was not as concerned with what the student learned, but how they were able to articulate what they learned. The lessons were all conducted in partnership by the science support specialist and each participating teacher. The researcher did visit the science classroom prior to each lesson to explain what was going to happen following the lesson, why they were doing the research, and asked if anyone had questions. Once the lesson started, the researcher walked across the hallway to the library and set up the area to conduct the interviews. The researcher waited in library for the science investigation to be completed. Following the lesson, one half of the students went back to their regular classroom with their participating teacher, and one half of the students stayed in the science classroom with the science support specialist. The science support specialist sent students to the library to be interviewed one at a time. The participating students were assigned their anonymous identification codes as they came to the interview with the researcher. The science support specialist and participating teachers were never given the students' codes. The post-lesson interview protocol was followed (see Appendix F). During the interview the researcher also actively observed the students and took field notes. Audio recordings and written documentation data were collected. While one student was being interviewed the remaining students in the "interview first" group engaged in free time under the supervision of the science support specialist in the science classroom. At the same time, the participating teacher was with the "teacher-driven assessment first" group and was administering the teacher-driven assessment to the other half of the students. On day two, in each class, the order of the assessment students took first was reversed.

The technology used for the audio recordings consisted of two digital audio recorders that were placed on the table in front of the student participant being interviewed. No students were videotaped or photographed during the research.

A duplicate copy of the teacher-driven assessment data was provided to the researcher. The participating classroom teachers retained the original documents for their records, and as part of their lesson planning. The story interview session data were uploaded for analysis from the digital audio recorders to a secure computer at the researcher's higher education institution.

Data analysis and interpretation. According to Creswell (2013), the first step when conducting a data analysis is to organize the data. He refers to this as "the first loop in the spiral" (p. 182). In the second step, the researcher will read and assess the data that has been collected. The third step in the data analysis is the process of describing, classifying, and interpreting the data into codes, categories, and themes. Creswell (2013) states that codes found during this phase can represent:

- Information that researchers expect to find before the study
- Surprising information that researchers did not expect to find
- Information that is conceptually interesting or unusual to researchers (and potentially to participants and audiences) (p. 186)

The researcher must take the raw data and assign it to emerging categories. Carefully studying and inspecting the data will allow useful patterns to emerge that the researcher can connect to broader issues and theories. The researcher must continue to compare the data until he/she reaches the point of theoretical saturation – where the addition of any additional data is not successful in contributing new material to the codes, categories, or concepts that have been presented (Denscombe, 2010).

In this project, the researcher analyzed three methods of data collection. The story interview results were transcribed and coded from the audiotaped sessions. The teacher-driven assessments were studied and coded. And the documented on-site observations were analyzed and coded. The researcher used a combination of analytic coding and constant comparison, using components of the grounded theory approach. More specifically, Creswell's (2013) template (p. 207) for coding narrative research was used as the basis for coding the story interview audiotaped sessions. Included in this template is a way to record key elements of the story interviews with the students. Those elements include:

- Chronology of the story – epiphanies and events
- Plot – characters, setting, problem, action, and resolution
- Three-dimensional space – interaction, continuity, and situation
- Themes – blank spaces for the emerging themes identified by the researcher

The Creswell template was altered to better fit this research (see Figure 1).

"Forming codes, or categories, represents the heart of qualitative data" (Creswell, 2013, p. 184). In addition to recording each student participant's interview data using Creswell's template, analytic coding was used as a progressive process of sorting and defining the whole data presented in the research. A coding scheme was developed to analyze the audiotaped sessions, the teacher-driven assessment scores, the data in the templates, and the researcher's field notes scene by scene and line by line. A framework for relational categories is displayed in the results section using data display tables.

Types of analytic coding. Three types of coding are utilized in the data analysis. Once the data is organized, an *open coding* process was used to identify concepts, behaviors – or any emerging similarities between participants and events. *Axial coding* was also used in the data analysis phase. Axial coding refers to coding that is done to make connections between existing categories or subcategories. And lastly, *selective coding* was used to integrate the categories obtained from the previous coding activity and build the developing theory (Pandit, 1996).

The trustworthiness of the analytic explanation was supported by using Hollway and Jefferson's (2000) four core questions for the interpretive process. Those questions are:

1. What do you notice?
2. Why do you notice what you notice?
3. How can you interpret what you notice?
4. How can you know that your interpretation is the "right" one?

This data analysis phase involved the separation of the data collected into its component parts. The study of complex things will allow the researcher to identify basic elements that assist him/her in identifying key elements of something that they wish to understand better (Denscombe, 2010).

Research Assumptions

The researcher assumed that the second-grade students understood the concept of telling a story about what they learned and that they understood the difference between fiction and non-fiction. The researcher also assumed that the students had enough language development to be able to use their learning story as a method of expression. It was also assumed that if the students chose to use props to assist them in communicating what they learned, that they had a relationship, or experience, with those props. Lastly, it is assumed that the second-grade student participants can read.

Chapter 4

Results

As evidenced in the literature review for this study, providing students with the best opportunity to reflect on, and articulate, what they learn in school is important. Teachers can struggle to discover what students *truly* know, or what/if they have learned following a lesson or inquiry. Knowing what students know is critical, not only for the student but for the teacher, so that learning can naturally progress, and the teacher can manage student learning successfully. This is especially important when students are on the cusp of learning, or in the zone of proximal development (ZPD). The ZPD is the area between what students can do and what they cannot do, or between what they know and do not know (Vygotsky, 1978). Vygotsky said that it is up to the teacher, or "more knowledgeable other" to provide support that advances each individual student's learning. This form of support is dependent on knowing where a student is in the learning process.

Given what we know about multiple intelligences and differentiated instruction, we have been woefully inadequate in *assessing* student learning by relying, primarily, on teacher-driven paper-and-pencil type assessments. The purpose of this study was to examine two contexts in which summative assessment allows students to articulate what they've learned following a science lesson. Do existing paper-and-pencil type (teacher-driven) assessments provide students with what they need to articulate their learning? Do alternative (post-lesson interview) summative methods help children articulate what they learned following a science lesson? This study proposes opening a dialogue related to the opportunities we give young students to express what they have learned by exploring another possible approach, and comparing and complimenting the results of an existing teacher-driven method.

In this chapter, the results of two different assessments are presented. Students were asked to engage in two types of assessments following two different science lessons – one a teacher-driven, paper-and-pencil assessment, and the other a post-lesson interview with the researcher.

This chapter will have three sections. The first section will introduce the study participants. The second section will share the teacher-driven assessment results, and the third section will share the results from the constant comparative coding method.

Meet the Participants

The teacher participants. The first participant to join the research project was a science support specialist in the MPS elementary building chosen as the site location. The science support specialist's highest level of education is a master's degree in educational administration. This participant also indicated they held another endorsement in science education. The science support specialist has been teaching for 15 years. The role of the science support specialist at this school, is to provide science lesson/inquiry opportunities to students and teachers twice a week (one hour each) in the science center, and then follow-up with science extension activities in the teachers' classrooms at various times throughout the week. The science support specialist does not have a role in assessing student learning. The classroom teachers manage and implement all assessment activities.

The second participant to join the study was a second-grade teacher from the same MPS as the science support specialist. Their highest level of education is a bachelor's degree in elementary and early childhood education, and they are certified to teach kindergarten through third grade as well as first through eighth grade. The teacher did not report any special endorsements, but has been teaching for 20 years. This teacher participant uses their state's model curriculum for a science framework along with the MPS's curriculum map. Additionally, this teacher uses a branded series of science videos along with a combination of demonstrations and lab experiences. For assessment, this teacher relies on paper-and-pencil assessments and uses a checklist in a rubric related to student performance.

The third participant to join the study was a second-grade teacher who is also from the same MPS as the science support specialist. The highest level of education that this teacher has achieved is a master's degree in early childhood education, and a reading endorsement. This participant has been teaching for 16 years. To teach science concepts this teacher uses a combination of information text, science journals, and labs in conjunction with prior background knowledge. This teacher reported using a combination of formative and summative assessment methods that includes observation, tests, and journaling.

The student participants. The students in this study are all enrolled in the same MPS, and were the current students of the science support specialist and the two participating teachers. Twenty-seven students had permission and were approved for the study. One student was absent on both days that data was collected reducing the total number of participants to 26. One student's family did not fill out the demographic data. All students at this MPS participate in the free and/or reduced lunch program. The youngest student in the study was seven years old and the oldest student was nine years old. The average age was eight years old. Of the twenty-five students who attended one or both data collection events, and returned demographic information, seventeen were male and eight were female. Seventeen students identified as black, three white, and five multi or bi-racial. Of the ten students who reported a preschool background, seven had attended Head Start, two a private child care/preschool, and one attended a public preschool. Twelve parent respondents were single parents, nine indicated both parents were in the household, and four indicated there was one biological parent with a step-parent (see Table 1 and Table 2). The participating teachers ranked the students as high, medium, or low-achieving students (see Table 3). The results were – eight high-achieving, nine medium-achieving, and nine low-achieving students.

Table 1
Family Demographic Data Classroom 1

Student (n=13*)	Age	Free/Reduced Lunch Program Y/N	Gender	Parent/ Guardian Status	Preschool Y/N	Race
SP1C1	7	Y	M	Both Parents	Y Head Start	Black
SP2C1	8	Y	F	Single Parent	Y Public Preschool	Black
SP3C1	8	Y	F	Both Parents	N	Black
SP4C1	Unknown	Y	M	Unknown	Unknown	Unknown
SP5C1	7	Y	M	Both Parents	Y Private Preschool	Multi-racial
SP6C1	8	Y	M	Parent and Step Parent	Y Head Start	Black
SP7C1	8	Y	M	Both Parents	N	Black
SP8C1	8	Y	M	Single Parent	Y Head Start	Black
SP9C1	8	Y	M	Both Parents	N	Black
SP10C1	9	Y	F	Single Parent	N	Black
SP11C1	9	Y	M	Single Parent	N	White
SP12C1	8	Y	M	Both Parents	N	Black
SP13C1	8	Y	M	Single Parent	N	Black

Table 2
Family Demographic Data Classroom 2

Student (n=13)	Age	Free/Reduced Lunch Program Y/N	Gender	Parent/ Guardian Status	Preschool Y/N	Race
SP1C2	9	Y	F	Single Parent	N	Black
SP2C2	8	Y	M	Single Parent	Y Head Start	Black
SP3C2	9	Y	M	Single Parent	N	Black
SP4C2	9	Y	F	Parent and Step Parent	Y Private Child Care	Multi-racial
SP5C2	8	Y	F	Both Parents	N	White
SP6C2	8	Y	M	Single Parent	Y Head Start and Public Preschool	Black
SP7C2	8	Y	M	Single Parent	N	Black
SP8C2	8	Y	F	Both Parents	Y Head Start	Black
SP9C2	9	Y	F	Parent and Step Parent	N	Multi-racial
SP10C2	8	Y	M	Both Parents	Y Head Start	Black
SP11C2	9	Y	F	Single Parent	N	White
SP12C2	8	Y	M	Single Parent	N	Multi-racial
SP13C2	7	Y	M	Parent and Step Parent	N	Multi-racial

Table 3
Participant Teachers' Student Achievement Rankings

Student (n=13)	**High Achieving**	**Medium Achieving**	**Low Achieving**	**Student (n=13)**	**High Achieving**	**Medium Achieving**	**Low Achieving**
SP1C1	**X**			**SP1C2**		**X**	
SP2C1			**X**	**SP2C2**			**X**
SP3C1		**X**		**SP3C2**		**X**	
SP4C1			**X**	**SP4C2**		**X**	
SP5C1		**X**		**SP5C2**	**X**		
SP6C1		**X**		**SP6C2**	**X**		
SP7C1		**X**		**SP7C2**	**X**		
SP8C1			**X**	**SP8C2**	**X**		
SP9C1	**X**			**SP9C2**			**X**
SP10C1	**X**			**SP10C2**		**X**	
SP11C1			**X**	**SP11C2**			**X**
SP12C1		**X**		**SP12C2**			**X**
SP13C1			**X**	**SP13C2**	**X**		

Teacher-Driven Assessment Results

Following each science lesson, both classroom teachers administered a teacher-driven assessment that was designed by the site's second grade science team. The science team consisted of the two second grade classroom teachers and the science support specialist taking part in this study. The assessments were part of the teachers' routine evaluative process to determine what students have learned following a science lesson. Each test, or assessment, included ten questions worth ten points each. The researcher had no input into the design or implementation of this routine assessment process. The only scores the researcher received were those of the students who were participating in this study (see Table 4). The study was designed to utilize the teachers' existing process in determining what was learned following a science lesson. The teachers reported that it takes one hour to design their paper-and-pencil assessments, ten minutes to distribute and explain the assessment, fifteen minutes for the students to take the assessment, five minutes to collect the assessments, and one hour to grade the them. The total time involved in developing and implementing the assessments is two hours and thirty minutes.

Following the first lesson (Energy), twenty-two students took the teacher-driven science assessment. Eight of those students scored above 70 percent. Fourteen students scored 70 percent or lower, and four students were absent for both the lesson and assessment. Following the second lesson (Solar System), twenty-four students took the teacher-driven science assessment. Sixteen student participants scored above 70 percent on the assessment. Eight students scored a 70 percent or lower, and two students were absent for both the lesson and assessment. The highest score on the teacher-driven assessment was 100 percent, and the lowest score was 20 percent.

There were two questions on the front page of the teacher-driven Energy assessment that had identical and nearly identical questions on the back side of the teacher-driven assessment. The first set of duplicate questions were – "Which of these has kinetic energy?" and, "If an arrow is flying through the air, it has what type of energy?" Between the two classrooms, thirteen students answered these questions differently, one correct and one incorrect. The second set of questions were identical, and both questions read – "Which of these has potential energy?" One question appeared on the front page of the teacher-driven assessment and one question appeared on the back page. Fourteen students answered the question correctly on one side of the assessment and then incorrectly on the other side of the assessment. In total, there were seventeen instances where a similar or exact question was answered both correctly and incorrectly by the same student participant.

Table 4
Teacher-Driven Assessment Results

Students Classroom One (n=13)	Energy Lesson	Solar System Lesson	Students Classroom Two (n=13)	Energy Lesson	Solar System Lesson
	%	%		%	%
SP1C1	100	N/A*	SP1C2	90	100
SP2C1	20	60	SP2C2	40	80
SP3C1	50	70	SP3C2	60	100
SP4C1	70	N/A*	SP4C2	70	70
SP5C1	60	90	SP5C2	90	100
SP6C1	80	80	SP6C2	70	100
SP7C1	60	100	SP7C2	40	80
SP8C1	40	60	SP8C2	80	100
SP9C1	100	100	SP9C2	90	100
SP10C1	90	100	SP10C2	50	70
SP11C1	N/A*	70	SP11C2	50	70
SP12C1	N/A*	90	SP12C2	70	100
SP13C1	N/A*	50	SP13C2	N/A*	100

*N/A = Absent

A question-by-question analysis. Each teacher-driven assessment had ten questions related to the concepts taught in each of two science lessons (see Appendix G and Appendix H). The teacher-driven assessment for the first lesson (Energy) was taken by 22 student participants from the two participating classrooms. The assessment had ten questions (see Table 5 and Table 6). The first question on the teacher-driven Energy assessment was, *"What is energy?"* Fifteen student participants answered this question correctly and seven student participants answered it incorrectly. Only one student participant addressed the definition of energy in the interview. The second question was, *"Which of these has kinetic energy?"* Eleven students answered this question correctly, and eleven students answered the question incorrectly. However, eight of the student participants answered a very similar question (#10) correctly on the same assessment. Additionally, eight student participants defined or demonstrated kinetic energy accurately during the interview with the researcher. Question three was, *"Which of these has potential energy?"* Thirteen student participants answered this question correctly and nine student participants provided an incorrect answer. Two of the nine students who answered the question incorrectly answered it a second time (#7) correctly. Four of the nine student participants were able to provide an accurate answer, related to this question, during the interview with the researcher. The fourth question was, *"Which of these is a source of mechanical energy?"* Twelve student participants answered this question correctly and ten students answered it incorrectly. Of those ten students, four of the student participants addressed the question accurately during the interview with the researcher. The fifth question was, *"Chemical energy is stored in _________?"* Seventeen student participants provided correct answers for this question. Five student participants answered this question incorrectly. Of the five students who provided a wrong answer, two provided an accurate answer related to chemical energy during the interview with the researcher. The sixth question on the teacher-driven Energy assessment was, *"What is the main source of light energy on Earth?"* Seventeen student participants answered this question correctly and five student participants answered this question incorrectly. Although one student talked extensively about heat energy, no student brought up the topic of light energy during the researcher's interview. The seventh question, *"Which of these has potential energy?"* was answered correctly by thirteen student participants and incorrectly by nine. However, three of the same students answered the same question correctly a previous time (#3) on the same assessment. In three cases, student participants addressed this question accurately during the interview with the researcher. The eighth question on the teacher-driven assessment was, *"Rubbing your hands together creates what type of energy?* Seventeen correct answers were provided for this question and five student participants provided a wrong answer. Of the five students who missed the question on the teacher-driven test, one talked about heat energy during the post-lesson interview. The ninth question on the teacher-driven assessment was, *"Electromagnet energy is energy from ________."* Eighteen student participants provided a correct answer for question nine and four student participants answered the question incorrectly. One of the four students did provide an accurate answer for this question during the interview with the researcher. The tenth, and final, question on the Energy assessment was, *"If an arrow is flying through the air, it has what type of energy?"* Fourteen student participants provided a correct answer for this question and eight student participants answered this question wrong. Of the eight, four answered the same question correctly the first time it was asked on the same assessment. Additionally, three students provided an accurate answer related to this question during the researcher's interview.

The teacher-driven assessment for the second lesson (Solar System) was taken by 24 students from the two participating classrooms. The assessment had ten, multiple-choice, questions (see Table 7 and Table 8). The first question on the Solar System teacher-driven assessment was, *"The sun is a ______."* A total of twenty-one students answered this correctly. Only three students from Classroom One answered this question incorrectly, and there were no wrong responses from Classroom Two. Of the three student participants in Classroom One that answered the question incorrectly, one provided the researcher with a correct answer related to the topic during the interview. The second question for this assessment was, *"How does the Sun help the Earth?"* All student participants in Classroom One answered the question correctly, and eleven from the second classroom also got it right. Only two student participants from Classroom Two answered it incorrectly. Of those two, one student answered the question correctly during the interview with the researcher. The third question was, *"Earth is a _________."* Twenty-two student participants provided a correct answer for this question. Two students from Classroom One answered the question incorrectly but no student from Classroom Two got it wrong. Of the two students from Classroom One who answered it wrong, one addressed it accurately during the interview with the researcher. The fourth question was, *"How many weeks does it take the moon to travel around the Earth?"* Twenty-one student participants were able to provide a correct answer for this question. Three student participants answered this question incorrectly and, of the three, one student was able to accurately answer this question during the interview with the researcher. The fifth question was, *"An axis is______."* Seventeen student participants answered this question correctly. Seven student participants answered the question incorrectly and no student participant addressed the topic of this question during the interview. Many student participants did demonstrate the Earth moving on its axis but did not say the term or explain the activity. The sixth question on the assessment was, *"What is the main source of light energy on Earth?"* Nearly all the student participants answered this question correctly (22 of 24). Only two student participants, one from each classroom, answered this question incorrectly. Neither student addressed it accurately during the interview with the researcher. The seventh question was, *"The Earth travels around the __________."* Twenty-one student participants chose the correct answer for this question. Three student participants answered this question incorrectly. Two out of those three students addressed this question accurately during the interview with the researcher. The eighth question on the teacher-driven assessment was, *"Astronauts have visited the __________."* Again, twenty-one student participants provided a correct answer for this question, and three student participants provided and incorrect answer. None of the three students addressed the topic during the interview. The ninth question on the teacher-driven assessment was, *"The moon changes as it goes through its _________."* Sixteen student participants gave correct answers for this question and eight student participants answered the question incorrectly. Of the eight, one did address the topic accurately during the post-lesson interview. Several of the student participants discussed the phases of the moon using alternative terms such as – life cycles, crest moons, and banana shaped moon. The tenth, and final, question for the Solar System lesson was, *"How many planets are in our solar system?"* Twenty-one students answered this question correctly on the assessment and three student participants answered the question incorrectly. No student addressed the question during the post-lesson interview with the researcher.

Table 5
Teacher-Driven Energy Assessment Questions and Results Classroom One

Question	Answer	Correct Responses (n = 10)
What is energy?	Energy is the ability to do work or cause change	6
Which of these has kinetic energy?	A ball that has been thrown and is moving through the air	7
Which of these has potential energy?	A basketball that is on a shelf about to fall	7
Which of these is a source of mechanical energy?	Wind	4
Chemical energy is stored in ______?	Batteries	8
What is the main source of light energy on Earth?	The sun	9
Which of these has potential energy?	A book sitting on the edge of a table about to fall	5
Rubbing your hands together creates what type of energy?	Heat	7
Electromagnet energy is energy from _____?	Electricity	8
If an arrow is flying through the air, it has what type of energy?	Kinetic	6

Table 6

Teacher-Driven Energy Assessment Questions and Results Classroom Two

Question	Answer	Correct Responses (n = 12)
What is energy?	Energy is the ability to do work or cause change	9
Which of these has kinetic energy?	A ball that has been thrown and is moving through the air	4
Which of these has potential energy?	A basketball that is on a shelf about to fall	6
Which of these is a source of mechanical energy?	Wind	8
Chemical energy is stored in ______?	Batteries	9
What is the main source of light energy on Earth?	The sun	8
Which of these has potential energy?	A book sitting on the edge of a table about to fall	8
Rubbing your hands together creates what type of energy?	Heat	10
Electromagnet energy is energy from _____?	Electricity	10
If an arrow is flying through the air, it has what type of energy?	Kinetic	8

Table 7

Teacher-Driven Solar System Assessment Questions and Results Classroom One

Question	Answer	Correct Responses (n = 11)
The sun is a _______?	Star	8
How does the sun help the earth?	It gives earth heat and light	11
Earth is a _________?	Planet	9
How many weeks does it take the moon to travel around the earth?	4 weeks	9
An axis is ________?	An imaginary straight line that something turns around	6
What is the main source of light energy on earth?	The sun	10
The earth travels around the ______?	Sun	9
Astronauts have visited the ______?	Moon	9
The moon changes as it goes through its _______?	Phases	6
How many planets are in our solar system?	8	10

Table 8
Teacher-Driven Solar System Assessment Questions and Results Classroom Two

Question	Answer	Correct Responses (n = 13)
The sun is a _______?	Star	13
How does the sun help the earth?	It gives earth heat and light	11
Earth is a _________?	Planet	13
How many weeks does it take the moon to travel around the earth?	4 weeks	12
An axis is ________?	An imaginary straight line that something turns around	11
What is the main source of light energy on earth?	The sun	12
The earth travels around the ______?	Sun	12
Astronauts have visited the ______?	Moon	12
The moon changes as it goes through its _______?	Phases	10
How many planets are in our solar system?	8	11

Constant Comparative Analysis

The post-lesson, student participant, interviews were conducted exclusively by the researcher. The audio recordings were transcribed, and the narratives from the recordings were coded using a constant comparative method (Glaser & Strauss, 1967; Corbin &

Strauss, 1990; Strauss & Corbin, 1994). The data were coded with numerous iterations, until the researcher reached saturation and no new (or relevant) codes emerged. An adapted version of Creswell's (2013) template (p. 207) for coding narrative research (see Figure 1) was used to organize the codes into categories and themes. The three main categories were – *Teacher-Driven Assessment, The Post-Lesson Interview, and Three-Dimensional Space*.

Teacher-driven assessment. Because the teacher-driven assessment and the post-lesson interview are so closely linked in this research, the teacher-driven assessments were automatically associated with the coding process. The first category in the constant comparative analysis was, *Teacher-Driven Assessment*. Three codes – here on out referred to as "themes," were identified. Those themes were:

- Perfect Scores
- Inconsistent Answers to Same Question
- Additional Gain Based on Interview

The first theme in this category was identified as, *Perfect Scores.* Two perfect teacher-driven assessment scores were seen in the Energy assessment. Eight students scored above average on the teacher-driven assessment following the energy lesson, and fourteen students scored below average. The range of scores was 20 percent to 100 percent. On the Solar System teacher-driven assessment, eleven student participants had a perfect score. Sixteen student participants scored above average and eight scored below average. The range of scores on the Solar System assessment was 50 percent to 100 percent. In total, there were 13 perfect teacher-driven assessment scores for 12 students. This represented 46 percent of the student participants in the study. Six of the students in this theme were ranked as high-achieving, three medium-achieving, and three low-achieving. In comparison, 77 percent of the total student participants were able to articulate an equivalent "perfect score" to the researcher during the post-lesson interview.

The second theme identified in this category was, *Inconsistent Answers to Same Question*. As mentioned in the Teacher-Driven Assessment Results section, two sets of questions on the Energy assessment were very similar or identical. The student participant responses to these questions were coded as inconsistent. Within the two classrooms, 14 student participants answered the same question both right and wrong for a total of 17 inconsistent responses.

The third theme that was presented during coding was, *Additional Gain Based on Interview*. Though this data will be discussed under a theme in the *Post-Lesson Interview* category, it also belongs to the *Teacher-Driven Assessment* category. While coding the interview narrative, it became apparent that many students who answered a question incorrectly on the teacher-driven assessment addressed the same question/concept with accuracy during their interview. Between the two assessments, this happened eighteen times for fourteen student participants. As a result, the researcher applied an *additional gain* to the teacher-driven assessment scores for the student participants who confirmed an understanding of the concepts/questions that they missed on the teacher-driven assessments. Table 9 highlights how combining the two methods of assessment would, in some cases, raise the teacher-driven assessment scores. The gain was lowest for high achieving students, raising their score an average of 10 points. For the medium and low-achieving students the gain was higher. Both medium-achieving and low-achieving student scores increased an average of 20 points. Overall, eight scores were increased by

10 points. Six scores increased 20 points. Three scores were increased 30 percent. And lastly, one score was increased by 40 points. Fourteen of the eighteen scores benefitting from the gain of a combined assessment would have resulted in a grade change from below average to above average.

Table 9
Gain from Post-Lesson Interview
Applied to Teacher-Driven Assessment Score

Student	Initial Energy Score Points Out of 100	Adjusted Energy Score Points Out of 100
SP3C1	50	80
SP4C1	70	90
SP5C1	60	70
SP8C1	40	60
SP10C1	90	100
SP3C2	60	80
SP4C2	70	90
SP5C2	90	100
SP6C2	70	80
SP8C2	80	90
SP10C2	50	90
Student	Initial Solar System Score Points Out of 100	Adjusted Solar System Score Points Out of 100
SP2C1	60	80
SP3C1	70	90
SP12C1	90	100
SP13C1	50	80
SP4C2	70	100
SP8C2	90	100
SP10C2	70	80

The post-lesson interviews. The second category identified in the constant comparative analysis was, *The Post-Lesson Interview.* The post-lesson student interviews spanned over two days, and were conducted one week apart. The average time of completion for the student participant interview was five minutes and 56 seconds. Each student participant was interviewed twice – one time following the Energy science lesson, and one time following the Solar System science lesson. A total of 46 post-lesson interviews were conducted over the two days. Each classroom alternated participating in the data collection in the morning and in the afternoon. On the first day of data collection, one half of the student participants in each class were interviewed by the researcher prior to taking the teacher-driven assessment. The other half of the students took the teacher-driven assessment prior to being interviewed by the researcher. On the second day, the order in which the student participants were interviewed and took the teacher-driven assessment were reversed.

Three types of documents were coded as part of this category. The first documents were the transcribed narratives from the interview audio-recordings. The second document coded was the researcher's anecdotal records from the interviews. And lastly, the third document coded was the anecdotal records concerned with each teacher-driven assessment question and its relationship with the post-lesson interview data. This data is referred to as the Question-By-Question Analysis and is found under the previous, *Teacher-Driven Assessment Results*, section of this chapter.

Seven themes emerged related to the *Post-Lesson Interview* category. The themes identified were:

- Demonstrated Enthusiasm

- Incongruent with Teacher-Driven
- Identified an Area of Interest
- Concepts vs. Terms
- Stated a Misconception
- Confirmed Knowledge of Most Concepts
- Talked Through to an Understanding

The first theme to discuss is, *Demonstrated Enthusiasm*. Based on the audio-recordings and the researcher's anecdotal records stemming from the post-lesson interview, an above average level of enthusiasm was apparent in 58 percent of the student participants. These students were noted as eager to talk during the interview, and displayed excitement using their tone of voice and/or their body language. In one instance a student participant was struggling to remember a learned concept, so the researcher reminded the student that it was okay to be finished talking. The student replied, "I don't want to be done. I want to continue this interview." Another student suggested that their "brain was going to explode" from all the information they learned. A trait these 15 students had in common was that they went into exceptional detail about what they learned in the science lesson/s. Of the fifteen students, five were identified by their teachers as high-achieving, six as medium-achieving, and four low-achieving. Thirty-three percent of student participants in the *Demonstrated Enthusiasm* group were considered high-achieving. Forty percent of the students were considered medium-achieving, and twenty-seven percent of the students were considered low-achieving. This particular theme did not seem to favor any type of student and supports the idea that all students can get excited about learning science.

The second theme to be identified in the *Post-Lesson Interview* category was, *Incongruent with Teacher-Driven*. This coding theme emerged shedding light on the fact that students performed inconsistently related to concepts/questions when comparing one type of assessments to the other. This theme has a direct relationship to the theme in the *Teacher-Driven Assessment* category – *Additional Gain Based on Interview*. While coding the interview narrative, it became apparent that many students who answered a question incorrectly on the teacher-driven assessment addressed the same question/concept with accuracy during their interview. Between the two assessments, this happened eighteen times for fourteen student participants.

The third theme found in the *Post-Lesson Interview* category was, *Identified an Area of Interest*. Fourteen, or 54 percent, of the student participants exhibited a strong interest in a topic associated with the science lesson – but not necessarily addressed during the lesson. Many of these student participants were the same students who were referred to in the *Demonstrated Enthusiasm* theme. The students were interested in the following science topics:

- Earth Topography
- How Moon Phases Develop
- Gravity (numerous students discussed)
- Heat and Humans
- Geckos
- Dwarf Planets (numerous students discussed)
- More about Day and Night
- Venus, in particular

- Reflection in Space
- The Perspective of the Solar System
- Temperature
- Rotation in Space

The fourth theme discovered in the *Post-Lesson Interview* category was, *Concepts vs. Terms*. During the post-lesson interview, six students articulated terms related to the science lessons but struggled with conveying an understanding of the concept/s. This was evident seven times in post-lesson interviews (one student fell into this group twice).

Also during the post-lesson interview, there were student participants who articulated the concept/s related to the science lesson but could not recall any (or accurate) terms related to the lesson. This theme involved eleven students a total of thirteen times. Two students fell into this category two times.

The fifth theme found in the *Post-Lesson Interview* category was, *Stated a Misconception*. During the post-lesson interview with the researcher, nine student participants discussed thirteen misconceptions related to the science lessons. The misconceptions discovered were:

- Kinetic energy can move if it wants to
- At night, the sun turns into stars
- There are two types of energy – slow and fast
- The chemical was making a noise – so it's a noise chemical
- The moon stops orbiting when the day comes up
- The only thing we need is heat and light
- The moon phases are life cycles

- It takes one month, or one minute, or sixty seconds, for the earth to orbit the sun
- Pluto is no longer a planet – turning into a star
- The moon *or* earth orbits around the sun (as in takes turns)
- The sun moves around to different places during the day
- If the moon didn't shine, the earth would flood
- There is an East Pole

Three of the participating students who articulated misconceptions were ranked as high-achieving, four were ranked as medium-achieving, and two were ranked as low-achieving. Students were not prompted to answer questions, in specific ways or, to point out misconceptions.

The sixth theme identified in the *Post-Lesson Interview* category was, *Confirmed Knowledge of most Concepts*. During the post-lesson interview, there were 20 students who the researcher felt articulated a full understanding of the concepts related to the science lesson/s. Of these 20 students, 8 articulated this understanding for both science lessons (Energy and Solar System). This means that 77 percent of the total students in this study were able to articulate the equivalent of a perfect score during a post-lesson interview – as compared to 46 percent of the total student participants who accomplished a perfect score on a teacher-driven assessment.

The seventh theme identified in the post-lesson interview coding process was, *Talked Through to an Understanding*. A total of nine times, seven student participants talked themselves through a fact/concept that they initially presented to the researcher with uncertainty or incorrectly. Here is one example of a student participant talking through the process:

> I learned about energy. You can like, with some stuff, you can, like move it with your hands. Or you can, umm, learn, umm. You can see how high it's going to fall. There are about ten energies but I can't say their names. (researcher encourages student to take their time) Umm, there's a energy that starts with a C. And that is the energy that doesn't move. I think. Oh, it's kinetic. That is like, when something is not moving. Like this (demonstrated with a toy car not moving). And there's potential. Potential is when…wait, I messed up! Kinetic is where you move something and potential is just, when, something is just sitting.

Of the student participants who were identified in this category, only one was ranked as a high-achieving student. One student participant was ranked as low-achieving with the remaining four, and majority, being ranked as medium-achieving.

Three-dimensional space. The third category identified in the constant comparative analysis was, *Three-Dimensional Space*. Three-dimensional space was used to categorize the physical interaction between the student participants and the space around them, as well as the props available to them during the interview. Two themes emerged within this category. The identified themes were:

- Used Props
- Communicated Kinesthetically

The first theme identified in the *Three-Dimensional Space* category was, *Used Props*. The researcher provided props for the student participants to use, if they chose, to demonstrate what they learned in the science lessons. These props consisted of items they could build a model with (e.g., Styrofoam balls to demonstrate planets orbiting), recreate concepts with (e.g., toy cars, balls), and items they could draw or write on (e.g., magnetic doodle pad, paper, markers). Between the two classrooms, 19 students used props in 27 post-lesson interviews with the researcher. Of those 19 students, two were ranked as high-achieving, eight as medium-achieving, and nine low-achieving. High-achieving students used props less often to demonstrate what they've learned. Only 11 percent of the student participants who used props were ranked high-achieving. Forty-two percent of the student participants who decided to use props to help them articulate what they learned were ranked as medium-achieving. Forty-seven percent of the students who used props to articulate what they learned were ranked as low-achieving students.

The second, and final, theme identified in the *Three-Dimensional Space* category was, *Communicated Kinesthetically*. Six student participants used a form of bodily movement to demonstrate what they learned in their science lesson/s. The students used their bodies get their point across when talking with the researcher. Two of the student participants were ranked high-achieving, three medium-achieving, and one was ranked low-achieving. These students engaged in the following forms of kinesthetic movement:

- Used only their arms and hands to draw the shape of the sun
- Rocked their body back like football players block one another to demonstrate kinetic energy
- Jumped up and down to demonstrate kinetic energy
- Pretended to be the sun in a large universe using their arms to become more circular

- Rotated around the interview table like the moon rotates around the earth
- Motioned arms and hands in various directions to point out where everything is in the universe

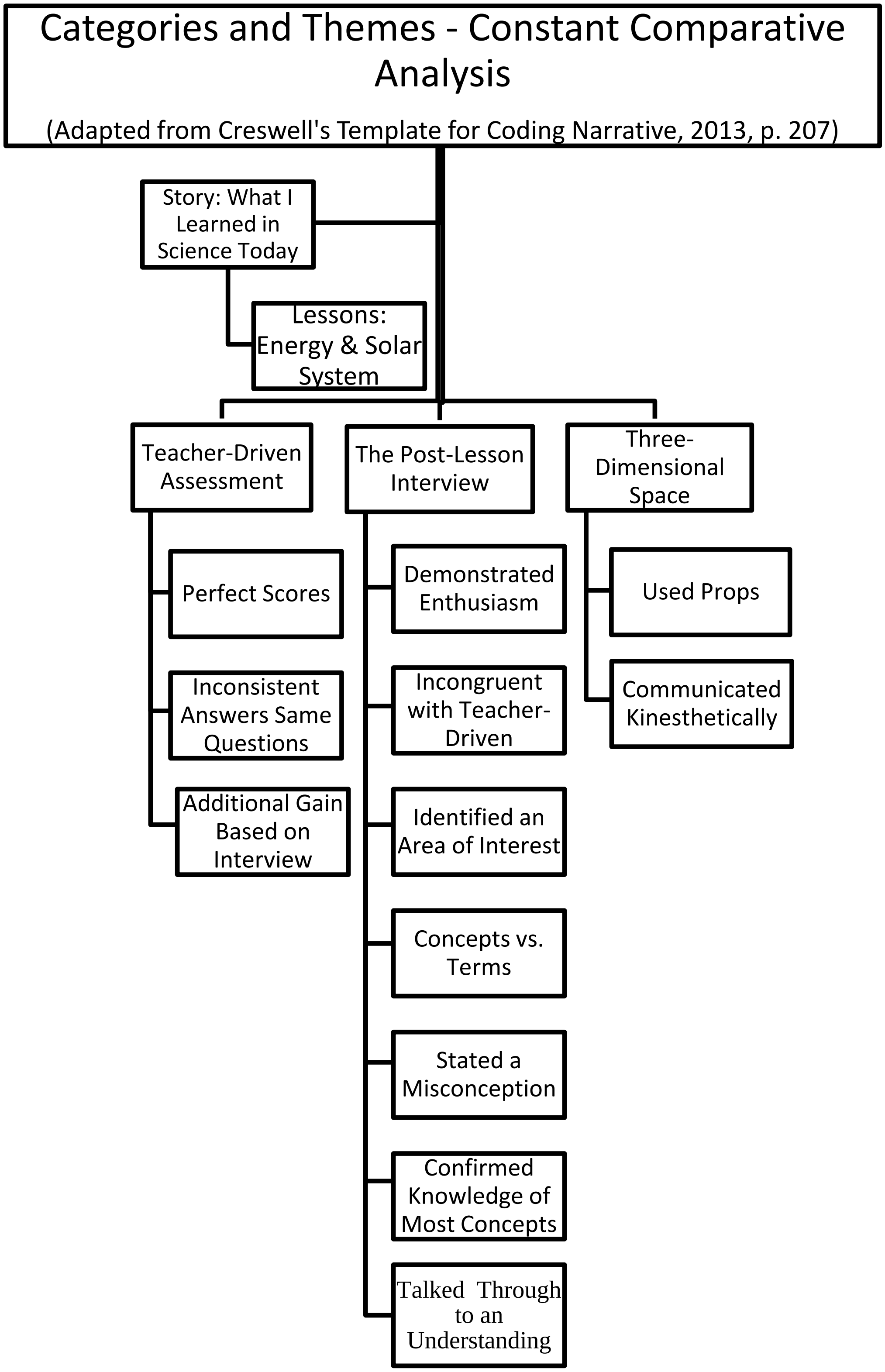
Categories and Themes - Constant Comparative Analysis
(Adapted from Creswell's Template for Coding Narrative, 2013, p. 207)
Story: What I Learned in Science Today
Lessons: Energy & Solar System
Teacher-Driven Assessment
The Post-Lesson Interview
Three-Dimensional Space
Perfect Scores
Inconsistent Answers Same Questions
Additional Gain Based on Interview
Demonstrated Enthusiasm
Incongruent with Teacher-Driven
Identified an Area of Interest
Concepts vs. Terms
Stated a Misconception
Confirmed Knowledge of Most Concepts
Talked Through to an Understanding
Used Props
Communicated Kinesthetically

Figure 1

Chapter 5

Individual Student Analyses

A Student-By-Student Narrative

In this chapter, a student-by-student narrative will highlight results from both the teacher-driven assessments and the post-lesson interviews. This narrative will consider the teacher-driven assessment results, the transcribed post-lesson interviews, the researcher's anecdotal records, and the themes that emerged during the coding process. Each student participant will have two narratives – one for each science lesson.

SP1C1 energy (interview assessment followed by teacher-driven assessment). SP1C1 was ranked a high-achieving student. The teacher-driven assessment score following this lesson was 100 percent. SP1C1 did well on the teacher-driven test and seemed to want to discuss additional items in the post-lesson interview. However, they were so excited about sharing what they learned that they had to be prompted to follow through with one topic before beginning a discussion on another. During the interview, the student participant brought up all the energies they learned about but was most interested in wind and mechanical. The student tended to stick to the terms and not go into much detail about the energy concepts. Toward the end of the interview, SP1C1 did state a misconception that energy can move "if it wants to." Although the student participant did not use props to demonstrate learning, they did engage in a form of kinesthetic movement to demonstrate the effects of light energy (how it shines with their arms and hands) on the earth.

SP1C1 solar system (absent).

SP2C1 energy (interview assessment followed by teacher-driven assessment). SP2C1 was ranked as a low-achieving student. The teacher-driven assessment score following this lesson was 20 percent. The questions that were answered incorrectly were:

1. "What is energy?"
2. "Which of these has potential energy?"
3. "Which of these is a source of mechanical energy?"
4. "Chemical energy is stored in _____."
5. "Which of the following has potential energy?"
6. "Rubbing your hands together creates what type of energy?"
7. "Electromagnet energy is energy from _____."
8. "If an arrow if flying through the air, it has what type of energy?"

Incorrect answers were:

1. "Light and sound are both energy"
2. "A football that has been thrown and is moving through the air"
3. "Food"
4. "Wind" and "Food" (circled two answers)
5. "A football flying through the air"
6. "Chemical"
7. "Sun"
8. "Potential"

SP2C1's first response when asked about what they learned in science today was, "Hmm, nothing!" When asked if they were sure they didn't learn anything, they picked up a ball and bounced it on the table. The student participant then picked up a car and demonstrated that when the car is moving ("like this") it is one type of energy and when the car isn't moving (held it still) it is another energy. It seemed to be difficult for SP2C1 to stayed focused on the interview. Though the student could not remember the names of the two opposite energies, it was clear that the student had learned the concept of the different energies. This student failed the teacher-driven assessment with an extremely low score. Though it wasn't a goal of this study to ask the same questions as the teacher-driven assessment in the post-lesson interview, it is the opinion of this researcher that this particular student would have benefitted greatly by doing so.

SP2C1 solar system (teacher-driven assessment followed by interview assessment). The teacher-driven assessment score following this lesson was 60 percent. The questions that were answered incorrectly were:

1. How many weeks does it take the moon to travel the earth?
2. An axis is _____.
3. Astronauts have visited the _____.
4. The moon changes as it goes through its _____.

Incorrect answers were:

1. 1 week
2. A planet
3. Mars
4. Day

SP2C1 successfully demonstrated the nature of orbiting planets, around the earth and sun, by using the available props. Although the student participant answered two of the abovementioned questions incorrectly on the teacher-driven assessment, they did convey the correct answers during the interview. SP2C1 also talked through incorrect ideas to a better understanding of the Solar System concepts. The student participant also demonstrated the tilt of earth's axis but was not able to define what an "axis" was on paper. In the post-lesson interview, this student was able to articulate an overall understanding of the concepts related to the Solar System lesson. If these things are taken into consideration, it would raise this student's teacher-driven assessment score from a 60 percent to an 80 percent.

SP3C1 energy (interview assessment followed by teacher-driven assessment). SP3C1 was ranked as a medium-achieving student. The teacher-driven assessment score following this lesson was 50 percent. The questions that were answered incorrectly were:

1. "What is energy?"
2. "Which of these has kinetic energy?"
3. "Which of these has potential energy?"
4. "Which of these is a source of mechanical energy?"
5. "Which of these has potential energy?"

Incorrect answers were:

1. "Energy is something that makes things jump"
2. "A ball on a shelf about to fall"
3. "A football that has been thrown and is moving through the air"
4. "Lights"
5. "A car rolling down a hill"

During the post-lesson interview, SP3C1 talked about three types of energy but was not able to name them. Through using props, it was evident that the student participant understood the concept of potential, kinetic, and chemical energy. On the teacher-driven assessment, SP3C1 answered three questions about potential and kinetic energy incorrectly. If credit for understanding the difference between these energies was applied to the teacher-driven assessment, the student's score would increase from 50 percent to 80 percent. SP3C1 did state a misconception that the sun turns into starts at night.

SP3C1 solar system (teacher-driven assessment followed by interview assessment). SP3C1 was ranked as a medium-achieving student. The teacher-driven assessment score following this lesson was 70 percent. The questions answered incorrectly were:

1. The sun is a _____.
2. An axis is a _____.
3. The moon changes as it goes through its _____.

Incorrect answers were:

1. Planet
2. Moving objects
3. Motions

Although SP3C1 missed the question addressing the term moon "phases," they brought up "phases" voluntarily during the post-lesson interview. Additionally, SP3C1 may have recognized that they called the sun a "planet" incorrectly because at one point they claimed that the sun turned from a "planet" to a "star" at night. Though the reason why they believe the term changes is unknown, it's clear that they have some understanding that the sun is a star. Because of this information it isn't unreasonable to believe that SP3C1's score on the teacher-driven assessment could have been raised from a 70 percent to a 90 percent. This student participant provided more detailed information during the post-lesson interview – such as naming the moon phases, attempting to name the planets, and demonstrating concepts using props. The student was less proficient with terms than concepts, but the researcher was confident this student had an overall understanding of the concepts shared in the Solar System lesson.

SP4C1 energy (interview assessment followed by teacher-driven assessment). SP4C1 was ranked as a low-achieving student. The teacher-driven assessment score following this lesson was 70 percent. The questions answered incorrectly were:

1. "Which of these has kinetic energy?"
2. "Which of these is a source of mechanical energy?"
3. "Which of these has potential energy?"

Incorrect answers were:

1. "A ball on a shelf about to fall"
2. "Candy"
3. "A basketball bouncing in the gym"

During the post-lesson interview, the energy types that SP4C1 wanted to discuss were "fast" and "slow" energies (a misconception). It could be that the student was referring to kinetic energy. Later, the student said there is also an energy that is not moving and used a ball to demonstrate a still object – referring to potential energy. Given time, the student participant did recall that wind energy was mechanical and talked about heat being energy as well. If you consider the concepts that SP4C1 clearly understood in the post-lesson interview, the score of the teacher-driven assessment would change from a 70 percent to 90 percent (and could be argued 100 percent if you believe that "fast" and "slow" energies represented kinetic). SP4C1 had a good overall understanding of the concepts shared in the Energy lesson.

SP4C1 solar system (absent).

SP5C1 energy (interview assessment followed by teacher-driven assessment). SP5C1 was ranked as a medium-achieving student. The teacher-driven assessment score following this lesson was 60 percent. The questions answered incorrectly were:

1. "Which of these is a source of mechanical energy?"
2. "What is the main source of light energy on earth?"
3. "Rubbing your hands together creates what type of energy?"
4. "If an arrow is flying through the air, it has what type of energy?"

Incorrect answers were:

1. "Lights"
2. "Street lights"
3. "Mechanical"
4. "Nuclear"

SP5C1 was able to name several types of energy they learned about in class, but it wasn't clear if they could match the description with the name in most cases. However, it was clear that the differences were understood – between kinetic and potential and between sun (light) energy and heat energy. The student participant answered one question about kinetic energy wrong on the teacher-driven assessment but clearly defined it during the post-lesson interview. The student also knew that the energy that is not moving started with a "P." If this student was given credit for the kinetic question alone (missed on teacher-driven assessment), their score would be raised from a 60 percent to a 70 percent. Although there was one misconception, SP5C1 had a good overall understanding of the concepts related to the Energy lesson.

SP5C1 solar system (teacher-driven assessment followed by interview assessment). SP5C1 was ranked as a medium-achieving student. The teacher-driven assessment score following this lesson was 90 percent. The question answered incorrectly was – "An axis is _____." The incorrect answer was – "A planet." It was difficult to keep SP5C1 on task and discussing the Solar System lesson. It was also clear that, although they did well on the teacher-driven assessment, the student participant did not understand the concept of orbit. The misconception was – "When the day comes out the moon stops, and then it's (earth) off the moon and on the sun." For both lessons, the student could perform well related to scientific terms but seemed less confident about the concepts being taught. The student participant also used props during both post-lesson interviews.

SP6C1 energy (teacher-driven assessment followed by interview assessment). SP6C1 was ranked as a medium-achieving learner. The teacher-driven assessment score following this lesson was 80 percent. The questions answered incorrectly were:

1. "What is energy?"
2. "Which of these is a source of mechanical energy?"

Incorrect answers were:

1. "Light and sound are both energy"
2. "Lights"

SP6C1 was ranked as a medium-achieving student and did quite well on the teacher-driven assessment. It was difficult to keep the student participant's attention for the post-lesson interview, but they did demonstrate enthusiasm for science. The student hinted that they understood magnetic ("magnet") energy, and demonstrated it with a prop. Other than being proficient in the terms related to the lesson, the only learning that could be confirmed during the interview was that a type of energy was present when something moved.

SP6C1 solar system (interview assessment followed by teacher-driven assessment). SP6C1 was ranked as a medium-achieving student. The teacher-driven assessment score following this lesson was 80 percent. The questions answered incorrectly were:

1. "Earth is a _____."
2. "How many weeks does it take the moon to travel around the earth?"

Incorrect answers were:

1. "Star"
2. "2 weeks"

The two questions SP6C1 missed did not present themselves in the post-lesson interview. It could be argued that the student participant provided more detail than allowed on the teacher-driven assessment related to night/day, and temperature, based on the rotation of the earth and its proximity to the sun. SP6C1 did not engage with the props in the second post-lesson interview.

SP7C1 energy (teacher-driven assessment followed by interview assessment). SP7C1 was ranked as a medium-achieving student. The teacher-driven assessment score following this lesson was 60 percent. The questions answered incorrectly were:

1. "Which of these has kinetic energy?"

2. "Chemical energy is stored in _____."
3. "Which of these has potential energy?"
4. "If an arrow if flying through the air, it has what type of energy?"

Incorrect answers were:

1. "A book that is sitting on a desk"
2. "Moving objects"
3. "A car rolling down a hill"
4. "Potential"

Based on the post-lesson interview, SP7C1 definitively articulated an understanding of the energy concepts missed related to the questions on the teacher-driven assessment. If given credit for this, the student participant's teacher-driven assessment score, combined with data from the interview, would be changed from 60 percent to 100 percent. During the post-lesson interview, SP7C1 included the use of props to demonstrate an overall understanding of the Energy concepts presented in the science lesson.

SP7C1 solar system (interview assessment followed by teacher-driven assessment). SP7C1 was ranked as a medium-achieving student. The teacher-driven assessment score following this lesson was 100 percent. This student participant scored perfectly on the teacher-driven assessment, and it could also be said that they provided more detail in the post-lesson interview. This student relied more on explaining the concepts vs. using the associated terms. SP7C1 was enthused about the lesson and offered impressive details about planets such as – *Venus is extremely hot, dwarf planets do not have moons, and that they are thinking about making Pluto a planet again.*

SP8C1 energy (teacher-driven assessment followed by interview assessment). SP8C1 was ranked as a low-achieving student. The teacher-driven assessment score following this lesson was 40 percent. The questions answered incorrectly were:

1. "What is energy?"
2. "Which of these is a source of mechanical energy?"
3. "Which of these has potential energy?"
4. "Rubbing your hands together creates what type of energy?"
5. "Electromagnet energy is energy from _____."
6. "If an arrow is flying through the air, it has what type of energy?

Incorrect answers were:

1. "Light and sound are both energy"
2. "Lights"
3. "A basketball bouncing in the gym"
4. "Chemical"
5. "Water"
6. "Potential"

SP8C1 understood that the sun gave us energy but could not remember what it was called (heat energy). This student chose to use props to help demonstrate what they learned in both lessons. If the Energy post-lesson interview results were combined with the teacher-driven assessment score, this student's overall score would increase from 40 percent to 60 percent. SP8C1 had a better idea of the concepts behind the Energy lesson than remembering the terms associated with the lesson.

SP8C1 solar system (interview assessment followed by teacher-driven assessment). SP8C1 was ranked as a low-achieving student. The teacher-driven assessment score following this lesson was 60 percent. The questions answered incorrectly were:

1. "An axis is _____."
2. "Astronauts have visited the _____."
3. "The moon changes as it goes through its _____."
4. "How many planets are in our solar system?"

Incorrect answers were:

1. "The name of a really good drummer"
2. "Mars"
3. "Monkeys"
4. "1"

There was no clear indication that SP8C1's post-lesson interview would have contributed to their teacher-driven assessment score. This student did not remember many terms related to the science lesson but understood the concepts – such as when they talked about the equator and demonstrated how the earth's axis works. Additionally, the student participant was enthusiastic about the lesson and enjoyed discussing the concepts of orbit, rotation, and gravity – and Geckos! SP8C1 was easily distracted until they started using the props to demonstrate what they learned. The student did demonstrate an overall understanding of the concepts presented in the Solar System lesson during the post-lesson interview.

SP9C1 energy (teacher-driven assessment followed by interview assessment). SP9C1 was ranked as a high-achieving student and was enthusiastic about what they learned in science. The teacher-driven assessment score following this lesson was 100 percent. SP9C1 discussed nearly all the energies and could define them accurately. The student participant did not mention electromagnet energy and referred to potential energy as "position" but other than that there were no significant differences between the post-lesson interview assessment and the teacher-driven assessment.

SP9C1 solar system (interview assessment followed by teacher-driven assessment). SP9C1 was ranked as a high-achieving student. The teacher-driven assessment score following this lesson was 100 percent. In this case, the student participant's post-lesson interview would not have contributed to the teacher-driven assessment score but the additional facts and detail that the student provided – such as talking about our "perspective" in relation to our position in the solar system, was remarkable. SP9C1 demonstrated proficiency in all concepts related to the science lesson.

SP10C1 energy (teacher-driven assessment followed by interview assessment). SP10C1 was ranked as a high-achieving student. The teacher-driven assessment score following this lesson was 90 percent. The question answered incorrectly was – "Which of these has potential energy?" The student participant chose the answer – "A spinning bike wheel." However, the student answered a question about potential energy correctly on the back side of the assessment. Take the duplicate question off the teacher-driven assessment and this student's score increases from 90 percent to a perfect 100 percent. During the post-lesson interview, the student participant knew the concepts of all the energies but could not recall some of the names – such as electromagnet and kinetic. This is the only student who recalled and presented a definition for energy itself. SP10C1 did not choose to use props to demonstrate their learning.

SP10C1 solar system (interview assessment followed by teacher-driven assessment). SP10C1 was ranked as a high-achieving student. The teacher-driven assessment score following this lesson was 100 percent. SP10C1 would not have scored better on the teacher-driven assessment using the results from the post-lesson interview but did go into more detail about the solar system during the post-lesson interview, such as – sharing how the moon isn't lit up, that it's the sun's light reflecting on the moon. SP10C1 did not choose to use props to demonstrate what they learned in the lessons.

SP11C1 energy (absent).

SP11C1 solar system (interview assessment followed by teacher-driven assessment). SP11C1 was ranked as a low-achieving student. The teacher-driven assessment score following this lesson was 70 percent. The questions answered incorrectly were:

1. "The sun is a _____."
2. "An axis is _____."
3. "The moon changes as it goes through its _____."

Incorrect answers were:

1. "Planet"
2. "Moving objects"
3. "Day"

Nothing changed from the post-lesson interview to the teacher-driven assessment for SP11C1, but the detail provided about planet orbit and rotation made it clear that they learned these concepts. And on the teacher-driven assessment it was clear that they were struggling with terms and not concepts. This student participants used props to aid in demonstrating what they learned in the Solar System science lesson.

SP12C1 energy (absent).

SP12C1 solar system (interview assessment followed by teacher-driven assessment). SP12C1 was ranked as a medium-achieving student. The teacher-driven assessment score following this lesson was 90 percent. The question SP12C1 answered incorrectly was – "The earth travels around the _____." The incorrect answer was – "Moon." SP12C1 used the Styrofoam balls to demonstrate that the earth travels around the sun. Giving them credit for understanding this would bring their score on the teacher-driven assessment to a perfect 100 percent. It is true that this student took the teacher-driven assessment prior to the post-lesson interview, but they also initially got the answer wrong in the interview as well. It was after demonstrating and talking through their thoughts that the student participant changed their mind and provided evidence of knowing the correct answer. The student participant also shared knowledge of concepts not on the test such as, gravity.

SP13C1 energy (absent).

SP13C1 solar system (interview assessment followed by teacher-driven assessment). SP13C1 was ranked as a low-achieving student. The teacher-driven assessment score following this lesson was 50 percent. The questions answered incorrectly were:

1. "The sun is a _____."
2. "Earth is a _____."

3. "What is the main source of light energy on earth?"
4. "The earth travels around the _____."
5. "The moon changes as it goes through its _____."

Incorrect answers were:

1. "Planet"
2. "Star"
3. "House lights"
4. "Moon"
5. "Motions"

Although SP13C1 answered questions relating to what the earth and sun are incorrectly on the teacher-driven assessment, it was clear in the post-lesson interview that they understood the sun was not a planet and the earth was a planet. The student participant even went on to share that the earth was the third planet from the sun. They also knew there were eight planets and used props to demonstrate the learned concepts related to the Solar System. This student answered the question about light energy incorrectly on the teacher-driven assessment but discussed the sun giving us light energy in the post-lesson interview. Giving the student credit for these three questions, the student's score on the teacher-driven assessment would increase from 50 percent to 80 percent.

SP1C2 energy (interview assessment followed by teacher-driven assessment). SP1C2 was ranked as a medium-achieving student. The teacher-driven assessment score following this lesson was 90 percent. SP1C2 answered the following question incorrectly

– “Which one of these is a source of mechanical energy?” SP1C2 chose the answer – “Lights.” The correct answer was – “Wind.”

SP1C2 discussed four types of energy during the interview and correctly defined each one. They included – potential, kinetic, chemical, and magnetic. The student missed mechanical energy on the teacher-driven assessment and during the interview eluded to a form of energy that they had forgotten –mechanical energy. On both assessments, mechanical energy proved to be an area of weakness. However, the student produced good, and consistent, results on both types of assessment. The student was finished with the interview once they could recall and recite four of the five energies. SP1C2 also used props and demonstrated enthusiasm about learning science facts.

SP1C2 solar system (teacher-driven assessment followed by interview assessment). SP1C2 was ranked as a medium-performing student. The teacher-driven assessment score following this lesson was 100 percent. Though SP1C2 didn’t share all the learned facts that were on the teacher-driven assessment, it was more apparent in the post-lesson interview that the student grasped the magnitude and working relationships within the solar system. For example – the student participant was amazed that it takes eight minutes for the sun to “get like warm, and like electricity, to earth.” Additionally, the student participant knew that there were “a whole lot of stars out there (more than there are grains of sand on all the beaches on earth).” SP1C2 also used props to demonstrate some of the concepts they learned.

SP2C2 energy (interview assessment followed by teacher-driven assessment). SP2C2 was ranked as a low-achieving student. The teacher-driven assessment score following this lesson was 40 percent. The questions answered incorrectly on the teacher-driven assessment were:

1. "Which of these has potential energy?"
2. "Chemical energy is stored in _____."
3. "What is the main source of light energy on Earth?"
4. "Which of the following has potential energy?"
5. "Rubbing your hands together creates what type of energy?"
6. "If an arrow is flying through the air, it has what type of energy?"

Incorrect answers were:

1. "A person running outside"
2. "Food"
3. "Street lights"
4. "A football flying through the air"
5. "Mechanical"
6. "Potential"

SP2C2 seemed shy and at times look down and not talk. However, in the beginning of the interview, the student talked about learning about potential energy and provided an accurate description and example of potential energy. On the teacher-driven test, two of the questions missed were related to potential energy. During the interview, the student said this about energy – "If you push something it goes and if you don't push it, it don't go. Potential. And some things use batteries." It was clear during the interview that the student understood potential energy, but missed three questions related to this type of energy on the teacher-driven assessment. This student had a better understanding of potential energy than what was reflected on the teacher-driven assessment. If the interview was used to augment this student's score, it would have gone from a failing 40 percent to a 70 percent. Although the teacher-driven assessment missed that the student had learned about a type of energy, the interview assessment provided them an opportunity to recall that particular energy. The interview assessment was challenging due to the shy demeanor of the student participant, but it was evident that this student was proficient with the concepts rather than the terms.

SP2C2 solar system (teacher-driven assessment followed by interview assessment). SP2C2 was ranked a low-achieving student. The teacher-driven assessment score following this lesson was 80 percent. The questions the student participant answered incorrectly were:

1. "An axis is _____."
2. "The moon changes as it goes through its _____."

Incorrect answers were:

1. "Moving objects"
2. "Motions"

During the post-lesson interview, SP2C2 used three Styrofoam balls to demonstrate how the sun stays still and how the moon "goes" around earth. The student participant was preoccupied with the researcher's recording devices, and worried that the battery would expire. Because of this, SP2C2 didn't seem to be able to recall much about the lesson. In terms of adding anything to the overall assessment, it should be noted that SP2C2 seemed to struggle with accurate terminology but had a good grasp on the overall concept of the solar system. This student talked through concepts they were struggling with, and used props to help articulate what was learned.

SP3C2 energy (interview assessment followed by teacher-driven assessment).

SP3C2 was ranked as a medium-achieving student. The teacher-driven assessment score following this lesson was 60 percent. The questions the student participant answered incorrectly were:

1. "What is energy?"
2. "Which of these has kinetic energy?"
3. "Rubbing your hands together creates what type of energy?"
4. "Electomagnet energy is energy from ______."

Incorrect answers were:

1. "Light and sound are both energy"
2. "A ball on a shelf about to fall"
3. "Chemical"
4. "Wind"

During the post-lesson interview, SP3C2 explained potential and kinetic energy with exceptional detail. Later in the interview the student also correctly defined chemical energy. The student remembered that there was another energy (magnetic), but did not proceed with a thorough definition (mentioned that "it moves"). The student did not mention mechanical energy but did get mechanical energy correct on the teacher-driven assessment. This student articulated in the interview that they understood the concept of kinetic and chemical energy although those questions were answered incorrectly on the teacher-driven assessment. If the score was adjusted to include the data in the interview the student's score would be raised from 60 percent to 80 percent. Based on the post-lesson interview, this student demonstrated verbally, and with props, that they learned both concepts and terms. Though the student missed the question – "What is energy?" It was also clear in the interview that the concept of energy was understood. The student chose the wrong answer of "wind" for the question – Electromagnet energy is energy from ______." However, on the teacher-driven assessment, the student understood that "wind" was mechanical energy. During the interview, the student had defined electromagnet energy as energy that moves. Technically, this wouldn't be wrong.

SP3C2 solar system (teacher-driven assessment followed by interview assessment). SP3C2 was identified as medium-achieving student. The teacher-driven assessment score following this lesson was 100 percent. Though some correct terms and facts, such as the number of planets that exist, did not present in the post-lesson interview, it was evident that this student participant had a true understanding of how the solar system works. Unlike this student's Energy post-lesson interview where they articulated concepts and terms, they were more proficient with concepts vs. terms in the Solar System post-lesson interview.

SP4C2 energy (interview assessment followed by teacher-driven assessment).

SP4C2 was ranked as a medium-achieving student. The teacher-driven assessment score following this lesson was 70 percent. The questions answered incorrectly were:

1. "Which of these has kinetic energy?"
2. "Which of these has potential energy?"
3. "Chemical energy is stored in _____."

Incorrect answers were:

1. “A ball on the floor not moving
2. “A person running outside”
3. Moving objects”

In the case of this question – “Which of these has potential energy?” SP4C2 answered the same question correctly later, on the same teacher-driven assessment. During the post-lesson interview, the student participant also defined potential energy correctly. The student mispronounced the names kinetic (“cognetic”) and potential (“prediction”), but articulated an understanding of both types of energy. SP4C2 used props to demonstrate knowledge and the only energy that the student participant did not mention, or define, in the post-lesson interview was mechanical. Using the same conceptual goals as the teacher-driven test, this student would have scored a 90 percent on the interview assessment – higher than the 70 percent on the teacher-driven assessment. SP4C2 stated one misconception, that the moon had *life cycles*. This student asked if the teacher-driven “test” was next. After confirming there was a test following the interview with the researcher, the student was happy and said that it’s easier to take a test after talking to someone about what they learned.

SP4C2 solar system (teacher-driven assessment followed by interview assessment). SP4C2 was ranked as a medium-achieving student. The teacher-driven assessment score following this lesson was 70 percent. The questions answered incorrectly were:

1. “How does the sun help the earth?”
2. “How many weeks does it take the moon the travel around the earth?”

3. "The moon changes as it goes through its _____."

Incorrect answers were:

1. "It keeps the moon in place"
2. "1 week"
3. "Day"

In the case of SP4C2, all incorrect answers on the teacher-driven assessment were recalled correctly in the post-lesson interview. The student also chose to use props to demonstrate some things they learned. Although the student participant did not mention a specific time that it takes the moon to travel around the earth, they did talk about and clearly understand that the moon traveled around the earth at a slow pace. Additionally, this student impressively named all the moons, categorizing them "life cycles." If the anecdotal data were combined with the teacher-driven assessment data, this student would have scored a perfect 100 percent.

SP5C2 energy (interview assessment followed by teacher-driven assessment).

SP5C2 was ranked as a high-achieving student. The teacher-driven assessment score following this lesson was 90 percent. The one question answered incorrectly on the teacher-driven assessment was – "Which of these is a source of mechanical energy?" The incorrect answer chosen was – "Lights." During the post-lesson interview, SP5C2 accurately defined three types of energy talked about in the lesson – kinetic, mechanical, and chemical. The student expressed difficulty remembering the names of the other types of energy but thought there were six. The student did accurately answer questions related to those three energies on the teacher-driven assessment. Likewise, the student provided a correct description of mechanical energy during the interview but missed the mechanical energy question on the teacher-driven assessment. The student was able to articulate an overall understanding of the Energy concepts in the post-lesson interview.

SP5C2 solar system (teacher-driven assessment followed by interview assessment). SP5C2 was ranked as a high-achieving student. The teacher-driven assessment score following this lesson was 100 percent. SP5C2 provided many quick facts about the solar system in the post-lesson interview but was less confident about the role of the planets together and individually. For example – the student first said that it takes one month for the earth to orbit the sun, and later said that it takes "one minute or sixty seconds" for the earth to orbit the sun. This student was more confident with the learned terms than with the concepts discussed in the Solar System lesson. SP5C2 did not use props during either post-lesson interview.

SP6C2 energy (interview assessment followed by teacher-driven assessment). SP6C2 was ranked as a high-achieving student. The teacher-driven assessment score following this lesson was 70 percent. The questions answered incorrectly were:

1. "Which of these has kinetic energy?"
2. Which of these is a source of mechanical energy?"
3. "What is the main source of light energy on Earth?"

Incorrect answers were:

1. "A ball on a shelf about to fall"
2. "Lights"
3. "Street lights"

The student incorrectly defined kinetic energy in the beginning of the interview, but later in the same interview provided a correct example of kinetic energy. Using the teacher-driven assessment results exclusively, it would be unclear if this student learned anything about kinetic energy. The student also remembered, and correctly defined, potential energy in the interview session. SP6C2 did not use props to demonstrate learning.

SP6C2 solar system (teacher-driven assessment followed by interview assessment). SP6C2 was ranked as a high-achieving student. The teacher-driven assessment score following this lesson was 100 percent. SP6C2 could recall and discuss most of the major concepts in the lesson. The student participant mentioned that they learned some "words" and understood that the moon travels around the earth and the earth travels around the sun. During the post-lesson interview, the student had a better understanding of the concepts and a more difficult time recalling the terms related to the lesson (although as mentioned above, the learned some words).

SP7C2 energy (interview assessment followed by teacher-driven assessment). SP7C2 was ranked as a high-achieving student. The teacher-driven assessment score following this lesson was 40 percent. The questions answered incorrectly were:

1. "What is energy?"
2. "Which is these has kinetic energy?"
3. "Which of these has potential energy?"
4. "Chemical energy is store in _____."
5. "What is the main source of light energy on Earth?"
6. "Which of these has potential energy?"

Incorrect answers were:

1. "Light and sound are both energy"

2. "A ball on a shelf about to fall"
3. "A person running outside"
4. "Moving objects"
5. "Street lights"
6. "A football flying through the air"

During the post-lesson interview assessment, SP7C2 was preoccupied with "trying to figure out how to be nice to people." The student participant was difficult to engage in an interview but did explain kinetic energy before stating he was finished. Though this student was ranked as a high achiever, the teacher-driven assessment grossly failed to reflect the student as a high achiever. The interview assessment could not be completed based on the emotional state of the participant

SP7C2 solar system (teacher-driven assessment followed by interview assessment). SP7C2 was ranked as a high-achiever. The teacher-driven assessment score following this lesson was 80 percent. The two questions answered incorrectly were:

1. "How does the sun help the earth?"
2. "What is the main source of light energy on earth?"

Incorrect answers were:

1. "It gives us seasons"
2. "Street lights"

In the post-lesson interview, SP7C2 did discuss the overall concept of orbit involving the earth, moon, and sun. However, when talking about length of orbit they suggested that either the moon or the earth orbits around the sun "every one month." Though this student participant answered this correctly on the teacher-driven assessment, they were more confused when trying to recall it on their own and did not choose to use props to help articulate what they learned.

SP8C2 energy (teacher-driven assessment followed by interview assessment). SP8C2 was ranked as a high-achieving student. The teacher-driven assessment score following that lesson was 80 percent. The two questions answered incorrectly were:

1. "Which of these has kinetic energy?"
2. "Which of these has mechanical energy?"

Incorrect answers were:

1. "A ball on the floor not moving"
2. "Lights"

During the post-lesson interview assessment SP8C2 recalled kinetic, potential, chemical, and light energy. The student also eluded to an understanding of electromagnet energy. The student explained something about each energy that confirmed they learned differences between them. It was apparent that the student understood kinetic energy even though answering one of two questions relating to kinetic energy was answered incorrectly on the teacher-driven test Information from the interview assessment and the teacher-driven assessment clearly indicates that this student understands kinetic energy. This would change their teacher-driven score from 80 to 90 percent. SP8C2 did use props in the post-lesson interview to help demonstrate what they learned about the phases of the moon and reflection.

SP8C2 solar system (interview assessment followed by teacher-driven assessment). SP8C2 was ranked as a high-achieving student. The teacher-driven assessment score following that lesson was 100 percent. Though SP8C2 would not have performed any better on the teacher-driven assessment based on the post-lesson interview for the Solar System lesson, the researcher did find that much more detail was provided in the interview – such as a keen understanding of how the moon phases are a result of the sun's shadow. This student used props during the post-lesson interview for Energy but chose not to during the Solar System post-lesson interview.

SP9C2 energy (teacher-driven assessment followed by interview assessment). SP9C2 was ranked as a low-achieving student. The teacher-driven assessment score following the Energy lesson was 90 percent. The only question answered incorrectly on the teacher-driven assessment was – "Which of the following has kinetic energy?" The incorrect answer was – "A ball on the floor not moving." This is the only energy that SP9C2 would discuss during the post-lesson interview assessment. If you count this accurate definition toward the student's teacher-driven assessment score it would raise it to a perfect score. The student did not seem interested in being interviewed so the session was ended.

SP9C2 solar system (interview assessment followed by teacher-driven assessment). SP9C2 was ranked a low-achieving student. The teacher-driven assessment score following that lesson was 100 percent. SP9C2 was more talkative in this post-lesson interview than in the first where they experienced "memory loss" and, in fact, was very enthusiastic. The student participant provided a nice demonstration with details about orbiting planets and moons, and understood how the earth, sun, and moon worked together to provide light and shadow on earth. Though the student did well on the teacher-driven test, the demonstration mentioned above indicates the student had a much deeper understanding of the relationship between the earth, sun, and moon than the teacher-driven assessment could demonstrate.

SP10C2 energy (teacher-driven assessment followed by interview assessment). SP10C2 was ranked a medium-achieving student. The teacher-driven assessment score following the that lesson was 50 percent. The questions answered incorrectly were:

1. "Which of these has kinetic energy?"
2. "Which of these has potential energy?"
3. "What is the main source of light energy on Earth?"
4. "Which of these has potential energy?"
5. "If an arrow is flying through the air, it has what type of energy?"

Incorrect answers were:

1. "A ball on a shelf about to fall"
2. "A football that has been thrown and is moving through the air"
3. "House lights"
4. "A car rolling down a hill"
5. "Mechanical"

During the post-lesson interview, it was confirmed that SP10C2 understood how potential and kinetic energies work and the difference between the two, though the two names were mixed up in the beginning of the interview. On the teacher-driven assessment, SP10C2 incorrectly answered all questions relating to kinetic or potential energy (4). Considering what knowledge was confirmed in the interview, the teacher-driven assessment would change from a 50 percent to a 90 percent. SP10C2 used props to help articulate what they learned in the Energy lesson.

SP10C2 solar system (interview assessment followed by teacher-driven assessment). SP10C2 was ranked as a medium-achieving student. The teacher-driven assessment score following that lesson was 70 percent. The questions answered incorrectly on the teacher-driven assessment were:

1. "The earth travels around the _____."
2. "Astronauts have visited the _____."
3. "How many planets are in our solar system?"

Incorrect answers were:

1. "Moon"
2. "Sun"
3. "10"

The second and third missed questions did not show up as learned knowledge in the post-lesson interview, but the first missed question ("The earth travels around the _____") was answered correctly using props during the conversation with the researcher. This knowledge added to the teacher-driven assessment would have raised SP10C2's score from a 70 percent to 80 percent.

SP11C2 energy (teacher-driven assessment followed by interview assessment).

SP11C2 was ranked as a low-achieving student. The teacher-driven assessment score following that lesson was 50 percent. The questions answered incorrectly were:

1. "What is energy?"

2. “Which of these has kinetic energy?”
3. “Which of these has potential energy?”
4. “Electromagnet energy is energy from _____.”
5. “If an arrow is flying through the air, it has what type of energy?”

Incorrect answers were:

1. “Energy is something that makes things jump”
2. “A book on a shelf about to fall”
3. “A person running outside”
4. “Wind”
5. “Potential”

SP11C2 seemed to be very shy. The student participant demonstrated kinetic energy using a prop (a ball) and their own body but could not remember the name of the energy. SP11C2 did recall that sun was light energy. This interview did not produce much conversation and was ended since the participant seemed uncomfortable and was not engaging in the conversation. SP11C2 was part of the group that answered two different questions related to the same energy both correctly and incorrectly on the teacher-driven assessment.

SP11C2 solar system (interview assessment followed by teacher-driven assessment). SP11C2 was ranked as a low-achieving student. The teacher-driven assessment score following that lesson was 70 percent. The questions answered incorrectly were:

1. “An axis is _____.”
2. “The moon changes as it goes through its _____.”

3. “How many planets are in our solar system?”

Incorrect answers were:

1. “A planet”
2. “Day”
3. “10”

Although SP11C2’s combined assessments would not have added anything to the teacher-driven assessment score, they were able to demonstrate a higher level of detail relating to the orbit of the earth and moon by physically demonstrating the activity. This student was shy during the Energy post-lesson interview but exhibited a high level of enthusiasm during the Solar System interview. Even though this student state a couple misconceptions, the researcher was confident that they had a good overall understanding of the concepts presented in the Solar System lesson.

SP12C2 energy (teacher-driven assessment followed by interview assessment).

SP12C2 was ranked as a low-achieving student. The teacher-driven assessment score for this lesson was 70 percent. The questions answered incorrectly were:

1. “Which of these has potential energy?”
2. “Which of these has potential energy?”
3. “If an arrow is flying through the air, it has what type of energy?”

Incorrect answers were:

1. “A spinning bike wheel”
2. “A football flying through the air”
3. “Potential”

The post-lesson interview did not produce much data related to what SP12C2 learned during the science lesson. The student participant talked about magnetic and heat energy but did not articulate the differences. This student exhibited very shy behavior and the researcher did not prolong the interview. On the teacher-driven assessment, the student participant answered a kinetic energy question correctly and one incorrectly. Since SP12C2 was so shy, it was impossible to confirm if the interview assessment could have verified additional learned knowledge.

SP12C2 solar system (interview assessment followed by teacher-driven assessment). SP12C2 was ranked as a low-achieving student. The teacher-driven assessment score for this lesson was 100 percent. This student was shy in the previous interview but seemed to enjoy the Solar System interview. Although SP12C2 had a perfect score on the second lesson, the level of detail provided in the post-lesson interview verified an even deeper understanding of light, as well as planets and rotation. This student did not use props during either post-lesson interview.

SP13C2 Energy (Absent).

SP13C2 solar system (interview assessment followed by teacher-driven assessment). SP13C2 was ranked as a high-achieving student. The score on this student participant's teacher-driven assessment was 100 percent. SP13C2 provided more detail in the post-lesson interview relating to some concepts. For example – facts about the earth's topography, and details about the phases of the moon. This student displayed enthusiasm but did not choose to use props to demonstrate what they learned.

Chapter 6

Discussion

This study compared two methods of summative assessment following a science lesson. Using the Reggio Emelia Approach and The Hundred Languages of Children as a theoretical framework, this study sought to determine if the use of interview, as an assessment tool, provides different results than existing teacher-driven (paper-and-pencil) tests. This chapter will provide conclusions, methodological reflections, limitations, implications, and recommendations.

Conclusions

Several themes were identified in this study. The main theme was that the post-lesson interview method of providing students with a way to articulate what they learned was complimentary to the teacher-driven assessment. The post-lesson interview exposed areas where the teacher-driven assessment suggested they had learned a concept, but the post-lesson interview confirmed they did not. Likewise, it exposed areas where they did appear to learn a concept but were not able to articulate it on the teacher-driven test. Additionally, when complimenting the teacher-driven assessment with the results from the post-lesson interview, 18 teacher-driven assessment scores were increased for a total of 14 student participants. Fourteen of the 18 scores changed a student participant's overall assessment score from below average to above average.

More student participants were able to articulate their proficiency related to the science topic during the post-lesson interview than on the teacher-driven assessment. Forty-six percent of student participants demonstrated proficient knowledge with a perfect score on at least one of two teacher-driven assessments. However, 77 percent of the student participants articulated proficiency on the topic during at least one of the two post-lesson interviews. Twenty-three percent of the student participants were not able to articulate what they learned in either of two post-lesson interviews. Fifty-five percent of the students who demonstrated proficiency on the teacher-driven assessment also articulated proficiency during one of the interviews. Forty-five percent of the student participants who articulated proficiency during an interview were not able to on a teacher-driven test. Only 21 percent of the student participants who demonstrated proficiency on a teacher-driven test, could not demonstrate proficiency during at least one of the post-lesson interviews.

The post-lesson interview, as a form of assessment, was also able to provide a wealth of data that the teacher-driven test did not. This data has the potential to not only better understand each individual learner, but to meaningfully inform lesson planning and curriculum development. The post-lesson interview assessment data provided the following discoveries.

Student enthusiasm. For the most part, a student cannot demonstrate enthusiasm on a paper-and-pencil test, nor is a teacher able to detect enthusiasm on this type of assessment. However, this study provided evidence that enthusiasm *can* be demonstrated by students during a post-lesson interview assessment. Fifty-eight percent of the student participants in this study demonstrated enthusiasm for the topics covered in the science investigation, as well as the investigation itself. These students were spread out evenly over the teacher rankings related to high, medium, and low achieving students.

Student interests. During the post-lesson interview the students identified a topic of interest related to, but not part of, the lesson. This type of data has the potential to provide unique feedback during the student assessment process. The data gathered from this type of feedback could also be used by the teacher to guide future lessons and learning opportunities. Additionally, this data can be used to “hook” students’ interest in the learning process. In this study, a topic of interest was presented by 54 percent of the student participants during the post-lesson interview.

An ability to talk through misconceptions. Students demonstrated nine times that additional learning could take place during the post-lesson interview assessment. Twenty-seven percent of the student participants in this study talked through a held misconception, or error in their thinking, and arrived at the correct answer on their own. These students were not prompted with additional questions. This event was most prevalent in the student participant group who were ranked by their teachers as medium-achieving students. Just as important, the post-lesson interview assessment was able to identify nine students who held 13 misconceptions that were not talked through to understanding. These misconceptions were not visible on the teacher-driven test. This evidence suggests that a post-lesson interview can play a key role in discovering what concepts are not fully understood. It was clear in this research that the post-lesson interview assessment served a dual purpose – data collector and teaching tool.

Kinesthetic communication. The post-lesson interview provided six students with a way to demonstrate their learning both verbally and kinesthetically. Bodily-Kinesthetic intelligence is one of Gardner’s multiple intelligences (2011) and it was evident that these student participants relied on bodily movement during their interview. It would not be possible for these students to demonstrate their knowledge using kinesthetic intelligence on the paper-and-pencil test.

Eager use of props. It is a well-known fact that children in the preschool age group benefit from using props to communicate their thinking. However, Salmon's (2001) research suggests that students between the ages of five and ten years old can also enhance their memory with the use of drawing or props. Students in this research were eager to use props. Nineteen of the participating students in the post-lesson interview assessment used props 27 times to communicate what they learned during the science investigation.

A relationship between assessment type and the ability to learn concepts. There was a relationship between students who knew concepts over terms and the assessment method. Eleven students knew concepts but could not recall the related terms 13 times in the post-lesson interview, and 10 of those 13 times the students scored below average on the corresponding teacher-driven assessment. This was a subtle finding but could provide important feedback to the teacher about the quality of the lesson and offer the possibility of additional/future learning opportunities.

Additionally, of the students who participated in the post-lesson interview, 90 percent were able to articulate useful knowledge related to the science investigation. The data in this research also suggested that even though many student participants benefitted from the multiple opportunities to convey their learning in the post-lesson interview, there were also students who expressed their proficiency on the teacher-driven assessment and were not interested in being interviewed. Though this group of students represented a minority of participants, it tells us that paper-and-pencil assessments may play an important role for some students to communicate what they've learned.

The post-lesson interview process provided students with multiple opportunities to express what they learned in multiple "languages." In addition, it potentially provides the teacher with immediate feedback on the quality of the lesson and opens the door to future and additional learning opportunities using topics of interest presented by students during the interview. The post-lesson interview was open-ended and, as a result, complimentary to the teacher-driven test. And for some students, the teacher-driven assessment was their preferred "language."

The teachers reported that it takes them two hours and thirty minutes to design, administer, and grade their science assessments. Based on the average interview time (5 minutes, 56 seconds), the researcher would be able to interview all student participants in one classroom (n = 13) in one hour and seventeen minutes. This suggests that interviewing students about what they learned is quicker than a teacher-driven paper-and-pencil assessment. Unfortunately, it is not known how much of the teacher assessment preparation and grading was classroom time vs. non-classroom time.

Fourteen students answered a similar or identical question inconsistently correct and incorrect on the teacher-driven assessment. If is unknown whether the science team designed their assessment this way intentionally, but it provided data that demands a deeper investigation into what appears to be inconsistent results in paper-and-pencil assessment methods. This appears to be a glaring deficiency related to this paper-and-pencil teacher-driven assessment. The question would be – did those 14 students pass or fail that question? Did they learn that concept or not? Does a one-word, multiple-choice, answer reflect all of what is going on inside the mind of a student? On the other hand, during the post-lesson interview, it was easy to recognize that learning took place based on what the students were able to articulate verbally or physically.

Methodological Reflections

It was the intention of the researcher to provide a story framework for the student participants to use as a way to articulate what they learned to the researcher during the post-lesson interview. Starting with the first interview, it was apparent that these second-grade students did not need a framework. They understood what an interview was, and how to engage in one effectively. The vast majority of the students were eager to "get interviewed." While the concept of story did not play out in its literal form, the students were eager to tell the "story" of what they learned during the science investigation.

The post-lesson interview was one open-ended question (Can you tell me the story about what you learned in science today?) because there was concern about leading the students to answers, and potentially affecting the teacher-driven assessment score. As a result, the interview wasn't directed against specific results that could be compared question-by-question with the teacher-driven assessment. Although this kept bias from entering the interview, it was frustrating not to be able to compare these two assessments in this way. For example – it could be said which students failed the teacher-driven assessment but not who failed the post-lesson interview. In this study, the interview was complimentary to the teacher-driven assessment rather than being truly comparable.

The researcher acknowledges that the number of participants in the study is low. However, it is a starting point. This type of research – which is not concerned with what students have learned on an assessment, but rather, how well the instrument enables them to articulate and/or demonstrate what they learned, is not common.

The quality of the science lessons and the paper-and-pencil teacher-driven assessments were not under the purview of the researcher. This was done to avoid any bias that could be entangled between them, as well as the post-lesson interview. Although the researcher maintains this opinion, some oversight over the quality and method of implementing these events would be worthwhile to think about in any future studies.

Limitations

Although the data clearly suggests that there is a more productive technique that can be used to provide young students with a way to articulate what they learned following a science lesson, there are limitations that should be discussed.

The most conspicuous limitation was the quality of the two science lessons. The design of the study purposefully facilitated the process of learning and assessing using the existing environment and procedures. This was done to ensure that participant student responses would be less likely to be influenced by the intrusion of the research study itself. And in fact, facilitating the process this way the study would also avoid bias that could flow between the lesson, the paper-and-pencil assessment, and the post-lesson interview. To avoid this intrusion, control over lesson quality was abandoned. This could also be a limitation for the study if the lessons did not maximize learning and the student participants had less data to express during the assessment methods. For the same reasons, the researcher relinquished control over the quality of the two teacher-driven assessments.

Another limitation was the personalities and moods of the student participants. Conducting interviews with students who exhibit shy personality traits, or who are not happy during the time of the interview, severely obstructed the data collection process.

However, one could also argue that the belief all students are unique learners and communicators is an overarching theme within this study. Previous research has established that all students do not learn in the same ways. Therefore, we should not expect all students to use the same "language" to articulate what they've learned.

Two things that didn't necessarily impact the study as much as the previously mentioned limitations were 1) the lack of relationship between the student and the researcher. Though the researcher visited the students prior to the onset of the data collection, the post-lesson interview results may have generated additional information if the student participants had the same relationship with the researcher as they had with the teacher participants. And lastly, 2) the subjective nature of the teacher student achievement rankings.

The study's number of participants will be addressed in the methodological reflections, but it should be mentioned in the limitations that there was no control over student participant attendance. In total, there were student absences during the data collection events. This significantly reduced the amount of data collected, as well as the ability to make comparisons at the student level and overall within the final data analysis.

Despite the limitations of a non-directed, open-ended post-lesson interview, it worked surprisingly well as a way for the participant students to recall and articulate much of what they learned. However, because it was so open-ended many student participants failed to describe some areas of the lesson. And because it was decided that leading the students to recall events associated with the lesson, the researcher was unable to fill in some gaps in what may have been learned.

Implications

This study clearly implies that most students are more successful demonstrating or articulating what they have learned following a science lesson in an interview, as opposed to engaging in a teacher-driven (paper-and-pencil) assessment. The post-lesson interview allowed students to use many "languages' to articulate what they learned in their science lesson/s. However, the research also pointed out that the teacher-driven paper-and-pencil assessment was the language of choice for some students and should remain a viable option for those students.

The study also implies that teacher-driven paper-and-pencil assessments can present incongruencies and inconsistencies that do not allow teachers to know definitively if a student has learned a concept related to the science lesson. Though it could not be proved that replacing a teacher-driven paper-and-pencil with an interview would work for all students, it was evident that a post-lesson interview could compliment the results of the type of teacher-driven assessment implemented in this study.

The study implies that a teacher/student interview type assessment can provide more useful information than a multiple-choice type paper-and-pencil assessment. An interview presents an opportunity for the students to use multiple languages to both demonstrate and facilitate their learning. It also allows them to be emotionally charged (e.g., enthusiastic) about their learning. It opens the door for more questions and for teachers to present additional topics of interests to interested students, as well as develop potential future topics for the entire classroom. A complete and accurate understanding of what students know is crucial when identifying learning needs and developing and implementing curriculum.

Recommendations

Future research should continue to examine and develop assessments that allow students to fully articulate what they've learned. Just as Gardner (2011) identified multiple intelligences of learning, we need to develop the multiple intelligences of expression – ways that students can tell us what they know. This study allowed students to tell us what they know using multiple methods that included traditional paper-and-pencil assessments and oral, and kinesthetic opportunities. While the limitations of the non-directed, and open-ended, interview did not allow the researcher to explore the full nature of the lesson the results of this limited interview were clearly superior to teacher-driven paper-and-pencil assessment. Future research should work to create and explore effective ways to allow students to express what they know in efficient and time-sensitive ways that can be integrated into the modern classroom.

References

Triangulation. (2015, September 10). Retrieved from Robert Wood Johnson Foundation: http://www.qualres.org/HomeTria-3692.html

(2008). *A guide to assessment in early childhood: Infancy to age eight.* Washington State Office of Superintendent of Public Instruction.

Abell, S. K., & Volkmann, M. J. (2006). *Seamless assessment in science: A guide for elementary and middle school teachers.* Portsmouth, NH: Heinemann.

Allen, A., Johnson-Whitt, E., Reinhart, M., North, T., Czerniak, C., Kaderavek, J., & Molitor, S. (2013). Making thinking visible in preschool science investigations. *Annual Meeting Mid-Western Educational Research Association*, (pp. 1-19). Evanston, IL.

Anderson, K. (2011). Science education and test-based accountability: Reviewing their relationship and exploring implications for future policy. *Science Education*, 104-129.

Antoniou, P., & James, M. (2014). Exploring formative assessment in primary classrooms: Developing a framework of action and strategies. *Educational Assessment, Evaluation, and Accountability*, 153-176.

Apgar, V. (1953). Proposal for a new method of evaluating the newborn infant. *Anethesia and Analgesia*, 260-267.

ASCD. (1993). *Curriculum update.* Retrieved from ASCD Learn, Teach, Lead: http://www.ascd.org/publications/curriculum-update/nov1993/teaching-young-children.aspx

Atkin, J. M., & Coffey, J. (2003). *Everyday assessment in the science classroom.* Arlington, VA: National Science Teachers Association Press.

Bagnato, S. J. (2007). *Authentic assessment for early childhood intervention.* New York: The Guilford Press.

Bagnato, S. J., Neisworth, J. T., & Pretti-Frontczak, K. (2010). *Linking authentic assessment and early childhood intervention: Best measures for best practices.* Baltimore: Paul H. Brookes Publishing.

Bagnato, S., & Ho, H. (2006). High-stakes testing with preschool children: Violation of professional standards for evidence-based practice in early childhood intervention. *KEDI Journal of Educational Policy*, 23-43.

Biddle, J. (2015, August 27). Primary grades in Reggio Emilia . (T. North, Interviewer)

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

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004, September). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan*, pp. 8-21.

Booth, D., & Barton, D. (2000). *Story works.* Ontario: Pembroke Publishers.

Bradley-Levine, J., & Carr, K. (2015). Critical theory and catholic social teaching: A research framework for catholic schools. *Journal of Catholic Education*, 26-43.

Brigance, A. H. (1985). Brigance Preschool Screen for Three-and Four-year-old Children. Curriculum Associates.

Bruner, J. (1991). The narrative construction of reality. *Critical Inquiry*, 1-21.

Bruner, J. (2004). Life as narrative. *Social Research*, 691-710.

Bruner, J. S. (2006). *In search of pedagogy. The selected works of jerome s. bruner.* London and New York: Routledge.

Bryman, A. (1988). *Quality and quantity in social research.* London: Unwin Hyman.

Cameron, H. (2005). Asking the tough questions: A guide to ethical practices in interviewing young children. *Early Child Development and Care*, 597-610.

Carr, M. (2001). *Assessment in early childhood settings, Learning stories.* Los Angeles: SAGE Publications, Inc.

Carr, M. (2011). Young children reflecting on their learning: Teachers' conversation strategies. *Early Years*, 257-270. Retrieved from Teaching and Learning Research Initiative: http://www.tlri.org.nz/sites/default/files/projects/9259_carr-summaryreport.pdf

Chang, N. (2005). Children's drawings: science inquiry and beyond. *Contemporary Issues in Early Childhood*, 104-106.

Charmaz, K. (2001). Grounded theory. In N. Denzin, & Y. Lincoln, *The American Tradition in Qualitative Research* (pp. 244-270). Thousand Oaks, CA: Sage Publications.

Christensen, P. (2004). Children's participation in ethnographic research: Issues of power and representation. *Children & Society*, 165-176.

Coleman, C., & Weidner, M. (2009, August 31). *Using stories to teach fluency strategies to young children* . Retrieved from Minnesota State University Mankato: https://www.mnsu.edu/comdis/isad12/papers/therapy12/coleman12.html

Collins, F. (1999). The use of traditional storytelling in education to the learning of literacy skills. *Early Child Development and Care*, 77-108.

Collins, F. (1999). The use of traditional storytelling in education to the use of literacy skills. *Early Child Development and Care*, 77-108.

Corbin, J., & Strauss, A. (1990). Grounded theory method: Procedures, canons, and evaluative criteria. *Qualitative Sociology*, 3-21.

Corbin, J., & Strauss, A. (2007). *Basics of qualitative research: Techniques and procedures for developing grounded theory.* Thousand Oaks, CA: Sage.

Council, N. R. (2008). *Early childhood assessment: What, what and how.* Washington D.C.: National Academies Press.

Crabtree, B. F., & Miller, W. L. (1992). *Doing qualitative research.* Newbury Park, CA: Sage.

Cress, S. (2004). Assessing standards in the real kindergarten classroom. *Early Childhood Education Journal*, 95-99.

Creswell, J. (2013). *Qualitative Inquiry & Research Design: Choosing Among Five Approaches.* Los Angeles: Sage.

Curenton, S. M. (2006). Oral storytelling: A cultural art that promotes school readiness. *YC Young Children*, 78.

Danby, S., Ewing, L., & Thorpe, K. (2004). The novice researcher: Interviewing young children. *Qualitative Inquiry*, 74-84.

Davis, P. (2007). Storytelling as a democratic approach to data collection: Interviewing children about reading. *Educational Research*, 169-184.

Delano, M., & Snell, M. E. (2006). The effects of social stories on the social engagement of children with autism. *Journal of Positive Behavior Interventions*, 29-42.

DeLuca, C., & Klinger, D. A. (2010). Assessment literacy development: Identifying gaps in teacher candidates' learning. *Assessment in education: Principles, policy & practice*, 419-438.

Denscombe, M. (2010). *The good research guide for small scale social research projects.* London: McGraw-Hill.

Denscombe, M. (2014). *The good research guide: for small-scale social research projects.* UK: McGraw-Hill.

Denzin, N. (1978). *Sociological methods.* New York: McGraw-Hill.

Denzin, N. (2006). *Sociological methods: A sourcebook.* Aldine Transaction.

Denzin, N. K., & Lincoln, Y. S. (2011). Introduction. In *The sage handbook of qualitative research* (pp. 1-19). Thousand Oaks, CA: Sage.

Dodge, D. T., Colker, L. J., & Heroman, C. (2001). A Teacher's Guide to Using The Creative Curriculum Developmental Continuum Assessment System. Teaching Strategies.

Early childhood curriculum, assessment, and program evaluation. (2003). Retrieved from National Association for the Education of Young Children: www.naeyc.org

Edwards, C. (2002). Three approaches from Europe: Waldorf, Montessori, and Reggio Emilia. *Early Childhood Research and Practice*.

Edwards, C. P., & Rinaldi, C. (2009). *The diary of Laura: Perspectives on a Reggio Emilia diary.* St. Paul, MN: Redlead Press.

Edwards, C., Gandini, L., & Forman, G. (1998). *The hundred languages of children: The reggio emilia approach - advanced reflections.* Greenwich, CT: JAI Press Ltd.

Edwards, C., Gandini, L., & Forman, G. (2012). *The hundred languages of children.* Santa Barbara: Praeger.

Eisner, E. (1975). The perceptive eye: Toward the reformulation of educational evaluation. *Occasional papers of the stanford evaluation consortium*. Stanford, CA: Stanford University.

Eisner, E. (2005). *Reimagining schools: The selected works of elliot eisner.* Abingdon, Oxon: Routledge.

Emde, R. N., Wolf, D. P., & Oppenheim, D. (2003). *Revealing the inner worlds of young children: The macarthur story stem battery and parent-child narratives.* New York: Oxford University Press.

Engel, S. (1995). *The stories children tell: Making sense of the narratives of childhood.* New York: W.H. Freeman and Company.

Enger, S. K., & Yager, R. E. (2009). *Assessing student understanding in science.* Thousand Oaks, CA: Corwin.

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

Ferrara, S., & DeMauro, G. E. (2006). Standardized assessment of individual achievement in K-12. In R. L. Brennan, *Educational measurement* (pp. 579-621). Westport, CT: Praeger Publishing.

Fiore, L., & Rosenquest, B. (2010). Shifting the culture of higher education: Influences on students, teachers, and pedagogy. *Theory Into Practice*, 14-20.

Fisher, J. (2004). *Starting from the child.* Buckingham: Open University Press.

Forman, G., & Fyfe, B. (1998). Negotiated learning through design, documentation, and discourse. In C. Edwards, L. Gandini, & G. Forman, *The hundred languages of children* (pp. 239-259). Greenwich, CT: Ablex Publishing Corporation.

Fyfe, B. (2012). The relationship between documentation and assessment. In C. Edwards, L. Gandini, & G. Forman, *The hundred languages of children* (pp. 273-294). Santa Barbara: PRAEGER.

Fyfe, B. (2012). The relationship between documentation and assesssment. In C. Edwards, L. Gandini, & G. Forman, *The hundred languages of children* (pp. 273-291). Santa Barbara, CA: Praeger.

Gardner, H. (2011). *Frames of mind.* New York: Basic Books.

Gardner, J., Harlen, W., Hayward, L., Stobart, G., & Montgomery, M. (2010). *Developing teacher assessment.* New York: McGraw-HIll.

Garrison, C., & Ehringhaus, M. (2007). *Formative and summative assessments in the classroom.* Retrieved from http://www.amle.org/Publications/WebExclusive/Assessment/tabid

Genishi, C., & Dyson, A. H. (2009). *Children, language and literacy.* New York: Teachers College Press.

Glaser, B. (1992). *Basics of grounded theory analysis: Emergence v. forcing.* Mill Valley, CA: The Sociology Press.

Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory: Strategies for qualitative research.* Chicago: Aldine.

Glazzard, J., Chadwick, Webster, A., & Percival, J. (2010). *Assessment for learning in the early years: Foundation stage.* Thousand Oaks, CA: SAGE Publications, Inc.

Glesne, C. (2011). *Becoming qualitative researchers.* Pearson: Boston.

Goulding, C. (1998). Grounded theory: The missing methodology on the interpretivist agenda. *Qualitative Market Research: An International Journal*, 50-57.

Grant, R., & Sugarman, R. (2004). Ethics in human subjects research: Do incentives matter? *Journal of Medicine and Philosophy*, 717-738.

Greenfield, D. (2015). Assessment in early childhood science education. In K. Trundle, & M. Saçkes, *Research in early childhood science education* (pp. 353-376). New York & London: Springer.

Greenfield, D. B. (2011). *Science instruction in preschool? Perspectives from the U.S.A. and why the time is right.* Cambridge: University of Cambridge.

Greenfield, D. B., Jirout, J., Dominguez, X., Maier, M. F., & Fuccillo, J. M. (2009). Science in the preschool classroom: A programmatic agenda to improve science readiness. *Early Education and Development*, 238-264.

Guddemi, M., & Case, B. (2004). *Assessing young children.* San Antonio: Pearson.

Gullo, D. F. (1994). *Understanding assessment and evaluation in early childhood education.* New York: Teachers College Press.

Gullo, D. F. (2004). *Understanding assessment and evaluation in early childhood education.* New York: Teachers College Press.

Harlen, W., & Qualter, A. (2009). *The teaching of science in primary schools.* London and New York: Routledge.

Harlen, W., & Qualter, A. (2014). *The teaching of science in primary schools.* New York: Routledge.

Harms, T., & Clifford, R. M. (1982). Assessing preschool environments with the early childhood environment rating scale. *Studies in Educational Evaluation*, 261-269.

Harris, M. (1968). *The rise of anthropological theory: A history of theories of culture.* New York: T.Y. Crowell.

Harrison, C., & Howard, S. (2011). Issues in primary assessment 3 How summative assessment sums up learning. *Primary Science*, 21-23.

Hattie, J. A. (2009). *Visible learning: A synthesis of over 800 meta-analyses related to achievement.* New York: Routledge.

Higgs, J. (2001). Charting standpoints in qualitative research. In H. Byrne-Armstrong, J. Higgs, & D. Horsfall, *Critical moments in qualitative research* (pp. 44-67). Boston: Butterworth-Heinemann.

High/Scope. (1992). High/Scope Child Observation Record for ages 21/2–6. Michigan: High/Scope.

Hollway, W., & Jefferson, T. (2000). *Doing qualitative research design differently: Free association, narrative, and the interview method.* Thousand Oaks, CA: Sage.

Hoover, N., & Abrams, L. (2013). Teachers' instructional use of summative student assessment data. *Applied Measurement in Education*, 219-231.

Irwin, D. M., & Bushnell, M. M. (1980). *Observational strategies for child study.* Holt, Rinehart, and Winston.

Isbell, R., Sobol, J., Lindauer, L., & Lowrance, A. (2004). The effects of storytelling and story reading on the oral language complexity and story comprehension of young children. *Early Childhood Education Journal*, 158-163.

Kaderavek, J., & Justice, L. M. (2002). Shared storybook reading as an intervention context: Practices and potential pitfalls. *American journal of speech-language pathology*, 395-405.

Kallio, A. A. (2015). Factional stories: Creating a methodological space for collaborative reflection and inquiry in music education research. *Research Studies in Music Education*, 3-20.

Kelley, M. F., & Surbeck, E. (1983). History of preschool assessment. In K. D. Paget, & B. A. Bracken, *The psychoeducational assessment of preschool children.* New York: Grune and Stratton.

Kincheloe, J., & McLaren, P. (2000). Rethinking critical theory in qualitative research. In N. Denzin, & Y. Lincoln, *Handbook of qualitative research* (pp. 279-313). Thousand Oaks, CA: Sage Publications.

King, P. H. (1975). *How do we know they're learning: Evaluation in the informal classroom.* Encino, CA: International Center for Educational Development.

Kincheloe, J., & McLaren, P. (2003). Rethinking critical theory and qualitative research. In N. K. Denzin, & Y. S. Lincoln, *The landscape of qualitative research* (pp. 433-488). Thousand Oaks, CA: Sage Publications.

Kortesluoma, R., Hentinen, M., & Nikkonen, M. (2003). Conducting a qualitative child interview: Methodological considerations. *Journal of Advanced Nursing*, 434-441.

Kostelnik, M., Soderman, A., & Whiren, A. (2011). *Developmentally appropriate curriculum.* Boston: Pearson.

Kuoch, H. (1998). *Social story interventions for young children with autism spectrum disorders: A comparative study* . Retrieved from The University of British Columbia: https://circle.ubc.ca/bitstream/id/27488/ubc_2001-0051.pdf

Lancy, D. F. (1993). *Qualitative research in education: An introduction to the major traditions.* New York: Longman.

Lash, M., Monobe, G., Kursun Koptur, D. U., & Black, F. V. (2016). Seven wonders of the early childhood world. *Childhood Education*, 236-246.

LeBuffe, P. A., & Naglieri, J. A. (2003). Devereux early childhood assessment—clinical form. Lewisville, NC: Kaplan Early Learning Company.

Lee, J., & Yoon, J. (2008). Teaching early childhood candidates how to assess children's inquiry skills in science learning. *Contemporary Issues in Early Childhood* , 265-269.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry.* Newbury Park, CA: Sage.

Looney, J. (2011). Integrating formative and summative assessment: Progress toward a seamless system? *OECD Education Working Papers, No 58,* 1-64. http://dx.doi.org/10.1787/5kghx3kbl734-en: OECD Publishing. Retrieved from OECD Education Working Papers, No. 58: http://dx.doi.org/10.1787/5kghx3kbl734-en

Lowenberg, J. S. (1993). Interpretive research methodology: Broadening the dialogue . *Advances in Nursing Science*, 57-69.

Luykx, A., Lee, O., Mahotiere, M., Lester, B., Hart, J., & Deaktor, R. (2007). Cultural and home language influences on children's responses to science assessments. *Teachers College Record*, 897-926.

Lyle, S. (2000). Narrative understanding Developing a theoretical context for understanding how children make meaning in classroom settings. *Journal of Curriculum Studies*, 45-63.

MacDonald, A. (2009). Drawing stories: The power of children's drawings to communicate the lived experience of starting school. *Australasian Journal of Early Childhood*, 40-49.

Mantovani, S. (2007). Italy: Pedagogy and curriculum. Infant toddler care. In R. S. New, & M. Cochran, *Early childhood education: An international encyclopedia, Vol. 4; The countries* (pp. 1113-1118, 1137-1141). Westport, CT: Praeger.

Marshall, C., & Rossman, G. (1999). *Designing qualitative research.* Thousand Oaks, CA: Sage.

Marshall, C., & Rossman, G. B. (2014). *Designing qualitative research* . Thousand Oaks, CA: Sage.

McDevitt, T., & Ormrod, J. (2010). *Child development and education.* Upper Saddle River, NJ: Merrill.

McNair, S., Bhargava, A., Adams, L., Edgerton, S., & Kypros, B. (2003). Teachers speak out on assessment practices. *Early Childhood Education Journal*, 23-31.

Meller, P., & Ohr, P. (2008). Family-oriented, culturally sensitive (FOCUS) assessment of young children. In L. Suzuki, & J. Ponterotto, *Handbook of multicultural assessment: Clinical, psychological, and educational applications* (pp. 565-593). San Francisco: John Wiley & Sons.

National Research Council. (2000). *How people learn: Brain, mind, experience, and school.* Washington, D.C.: National Academy Press.

National Research Council. (2001). *Classroom assessment and the national science education standards.* Washington D.C.: National Academies Press.

National Research Council. (2008). *Early childhood assessment: Why, what, and how.* Washington D.C.: The National Academies Press.

National Research Council. (2014). *Developing assessments for the next generation science standards.* Washington D.C.: National Academies Press.

NCWD/Youth. (2015, July 8). *Federation for Children with Special Needs.* Retrieved from National Collaborative on Workforce and Disability: http://fcsn.org/pti/topics/transition/tools/assessment.pdf

New, R. S. (2003). Reggio Emilia: New ways to think about schooling. *Educational Leadership*, 34-38.

Newborg, J., R, S. J., Wnek, L., Guidubaldi, J., Svincki, J., Dickson, J., & Markley, A. (1988). Newborg, J., Stock, J. R., Wnek, L., Guidubaldi, J., Svincki, J., Dickson, J., & Markley, A. . *Battelle Developmental Inventory with recalibrated technical data and norms: Screening test examiner's manual*. Allen, TX: DLM.

Oláh, L., Lawrence, N., & Riggan, M. (2010). Learning to learn from benchmark assessment data: How teachers analyze results. *Peabody Journal of Education*, 226-245.

Ollerenshaw, J. A., & Creswell, J. W. (2002). Narrative research: A comparison of two restorying data analysis approaches. *Qualitative Inquiry*, 329-347.

Onwuegbuzie, A. J., Leech, N. L., & Collins, K. M. (20081). Interviewing the interpretive researcher: A method for addressing the crises of representation, legitimation, and praxix. *International Journal of Qualitative Methods*, 1-17.

Ormrod, J. (2011). *Educational psychology*. Boston: Pearson.

Ornstein, A., & Hunkins, F. (2009). *Curriculum: Foundations, principles, and issues.* Boston: Pearson.

Pandit, N. (1996). The creation of theory: A recent application of the grounded theory method. *The qualitative report*.

Patton, M. (2002). *Qualitative Research and Evaluation Methods.* Thousand Oaks, CA: Sage Publications.

Pearson. (2005). *Achieving student progress with scientifically based formative assessment: A white paper.* Retrieved from Pearson: http://www.pearsoned.com/RESRPTS_FOR_POSTING/PASeries_RESEARCH/PA1.%20Scientific_Basis_PASeries%206.05.pdf

Plummer, K. (1983). *Documents of life: An introduction to the problems and literature of a humanistic method.* London: George Allen & Unwin.

Popham, W. J. (2008). *Transformative assessment.* Alexandria: ASCD.

Qualter, A. (2009). *The teaching of science in primary schools.* Routledge: New York.

Rimm-Kaufman, S. E., La Paro, K. M., Downer, J. T., & Pianta, R. C. (2005). The contribution of classroom setting and quality of instruction to children's behavior in kindergarten classrooms. *The Elementary School Journal*, 377-394.

Rinaldi, C. (2006). *In dialogue with Reggio Emilia: Listending, researching and learning.* London: Routledge.

Rinaldi, C. (2012). The pedagogy of listening: The listening perspective from reggio emilia. In C. Edwards, L. Gandini, & G. Forman, *The Hundred Languages of Children* (pp. 233-246). Santa Barbara: PRAEGER.

Rinaldi, C. (2013, June). Re-imaging childhood: The inspiration of Reggio Emilia education principles in South Australia. Adelaide, South Australia, . Retrieved from Adelaide Thinkers in Residence: http://thinkers.sa.gov.au/Thinkers/CarlaRinaldi/p513.aspx

Ritchie, J., & Lewis, J. (2003). *Qualitative Research Practice.* London: Sage.

Robert Wood Johnson Foundation. (2016, October 9). *Grounded Theory*. Retrieved from Robert Wood Johnson Foundation: http://www.qualres.org/HomeGrou-3589.html

Rutten, K., & Soetaert, R. (2012). Narrative and rhetorical approaches to problems of education Jerome bruner and kenneth burke revisited. *Studies in Philosophy and Education*, 327-343.

S, B., & Yeh Ho, H. (2006). High stakes testing with preschool children: Violation of professional standards for evidence based practice in early childhood intervention. *Journal of Educational Policy*, 23-43.

Sackes, M., Trundle, K., & Flevares, L. (2009). Using children's literature to teach standard-based science concepts in early years. *Early Childhood Education Journal*, 415-422.

Salmon, K. (2001). Remembering and reporting by children: The influence of cues and props. *Clinical Psychology Review*, 267-300.

Schneider, M. C., & Randel, B. (2010). Research on characteristics of effective professional development programs for enhancing educators' skills in formative assessment. In H. Andrade, & G. Cizek, *Handbook of formative assessment* (pp. 251-276). New York: Routledge.

Schwandt, T. (2007). *The SAGE dictionary of qualitative inquiry.* Thousand Oaks, CA: Sage.

Schwandt, T. A. (1997). *Qualitative inquiry: A dictionary of terms.* Thousand Oaks, CA: Sage.

Sciences, N. A. (1996). *National science education standards.* Washington D.C.: National Academy Press.

Seidman, I. (2006). *Interviewing as qualitative research: A guide for researchers in education and the social sciences.* New York: Teacher's College Press.

Shepard, L. (2006). Classroom assessment. In R. L. Brennan, *Educational measurement* (pp. 623-646). Westport, CT: Praeger Publishers.

Shmoop Editorial Team. (2008, November 11). *Poststructuralism basics.* Retrieved from Shmoop: http://www.shmoop.com/poststructuralism/the-basics.html

Snow, C. E. (2001). Knowing what we know: Children, teachers, researchers. *Educational Researcher*, 3-9.

Snow, C. V. (2008). *Early childhood assessment: Why, what, and how.* Washington D.C.: National Research Council of the National Academies.

Squires, J., Bricker, D., & Potter, L. (1997). Revision of a parent-completed developmental screening tool: Ages and Stages Questionnaires. *Journal of Pediatric Psychology*, 313-328.

Stiggins, R. J. (2002). Assessment crisis: The absence of assessment FOR learning. *Phi Delta Kappan*, 758-765.

Strauss, A. (1991). *Qualitative analysis for social scientists.* Cambridge: Cambridge University Press.

Strauss, A., & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques.* London: Sage Publications.

Strauss, A., & Corbin, J. (1994). Grounded theory methodology: An overview . In N. Denzin, & Y. Lincoln, *Handbook of qualitative research.* Newbury Park, CA: Sage Publications.

Supovitz, J. (2012). Getting at student understanding: The key to teachers' use of test data. *Teachers College Record*, 1-29.

Symonds, P. M. (1927). *Measurement in secondary education.* New York: Macmillan.

Tesch, R. (1990). *Qualitative research: Analysis types and software tools.* Bristol, PA: Falmer Press.

Theall, M., & Franklin, J. (2010). Assessing teaching practices and effectiveness for formative purposes. In K. Gillespie, & D. Robertson, *A guide to faculty development* (pp. 151-168). San Francisco: Jossey-Bass.

Thomas, J. (1993). *Doing critical ethnography.* Newbury Park, CA: Sage Publications.

Thorndike, E. L. (1913). *Introduction to the theory of mental and social measurements.* New York: Teachers College.

Trice, A. D. (2000). *A handbook of classroom assessment.* New York: Longman.

Usher, R. (1996). A critique of the neglected epistemological assumptions of educational research. In D. Scott, & R. Usher, *Understanding Educational Research* (p. 15). New York: Routledge.

Varelas, M., Pappas, C., Tucker-Raymond, E., Kane, J., Hankes, J., Ortiz, I., & Keblawe-Shamah, N. (2010). Drama activities as ideational resources for primary-grade children in urban science classrooms. *Journal of Research in Science Teaching*, 302-325.

Vygotsky, L. (1978). *Mind in society.* Cambridge: Harvard University Press.

Weber, E. (1984). *Ideas influencing early childhood education: A theoretical analysis.* New York: Teachers College Press.

Wertsch, J. V. (2002). *Voices of collective remembering.* Cambridge: Cambridge University Press.

Wilmot, A. (2005). *Designing the sampling strategies for social qualitative research: With particular reference to the office of national statistics' qualitative respondent register.* Retrieved from Office of National Statistics, UK: http://wwwn.cdc.gov/qbank/QUest/2005/Paper23.pdf

Wiltshire, M. (2013). *Understanding the HighScope approach: Early years education in practice* . New York: Routledge.

Wolcott, H. F. (1992). Posturig in qualitative research. In M. D. LeCompte, W. L. Millroy, & J. Preissle, *The handbook of qualitative research in education* (pp. 3-52). San Diego: Academic Press.

Wolcott, H. F. (2008). *Writing up qualitative research.* Thousand Oaks, CA: Sage.

Worth, K. (2010). Science in early childhood classrooms: Content and process. *STEM in early education and development.* Early Childhood Research and Practice.

Wortham, S. C. (1990). *Tests and measurement in early childhood education.* Columbus, OH: Merrill Publishing Company.

Wortham, S. C. (2011). *Assessment in early childhood education.* Pearson.

Yin, R. K. (2009). *Case study research: Design and method.* Thousand Oaks, CA: Sage.

Appendix A

Peer Debriefing Report

Peer Review and Debriefing

The peer reviewer was recommended, and considered a non-bias, disconnected peer, by the researcher's committee chair. The following report was provided to the researcher upon review of the methodology and data analysis sections of the dissertation:

> After reading the manuscript, I do not believe the researcher portrayed any preconceived hypotheses prior to collecting data. The researcher used constant comparative analysis to evaluate the results from both the teacher driven assessments and the post-lesson interviews. The researcher categorized responses in an appropriate manner.
>
> The researcher conducted the post-lesson interview with fidelity, while also being responsive to the student participants and using developmentally appropriate practices. I conducted a random reliability check on six students' responses (taken from transcription records) in both assessment tools. The results for these six, randomly chosen, students were accurately reported in the manuscript. The random reliability checks included the following student transcriptions:

SP7C2

SP11C1

SP2C1

SP2C2

SP9C2

SP1C1

These results were translated into empirical data. For example, percentages were used to show how many students received "perfect scores" or were highly interested in the subject matter. The research questions were supported by the data collection tools, participant selection, setting, and theoretical framework.

Appendix B

Parent/Guardian Demographic Questionnaire Form

Student Participant (SP) Name_____________________________

1) Participating student's gender (circle one): *M or F*
2) Participating student's race (circle one): *White – Black – American Indian – Alaskan Native – Asian – Native Hawaiian – Pacific Islander – Multi-racial – Other* (please specify)

   ____
3) Participating student's age: _______
4) Participate in Free and/or Reduced Lunch Program (circle one):

   Y or N
5) Did your son/daughter attend a preschool program (circle one):

   Y or N (circle one), and if so please indicate what type of experience he/she had (circle one):

   Public Preschool *Head Start* *Private Preschool*

   Fulltime Daycare
6) Family structure:

   Both Parents *Single Parent* *Parent & Step Parent*

   Grandparent *Other* (please specify):____

Appendix C

Teachers' Student Achievement Ranking Form

Student Achievement Rankings							
Student	High Achieving	Medium Achieving	Low Achieving	Student	High Achieving	Medium Achieving	Low Achieving
SP1C1				SP1C2			
SP2C1				SP2C2			
SP3C1				SP3C2			
SP4C1				SP4C2			
SP5C1				SP5C2			
SP6C1				SP6C2			
SP7C1				SP7C2			
SP8C1				SP8C2			
SP9C1				SP9C2			
SP10C1				SP10C2			
SP11C1				SP11C2			
SP12C1				SP12C2			
SP13C1				SP13C2			
SP14C1				SP14C2			

Appendix D

Participating Teacher Interview Form

Teacher Participant: ____

1) Highest Education Level__________________________________

2) Major____

3) Certification (grade)______________________________________

4) Special Endorsement/s____________________________________

5) Number of years teaching _________________________________

6) Gender: M or F

7) Race: *White – Black – American Indian – Alaskan Native – Asian – Native Hawaiian – Pacific Islander – Multi-racial – Other* (please specify) _________________________

8) How is teacher participant currently teaching science?

____

____

____

9) How is teacher participant currently assessing science?

____

____

____

Appendix E

List of Props Used in Post-Lesson Interview Assessment

Props Available to Students in Post-Lesson Interview Assessment

Energy Interview Props

Small toy cars

Small rubber balls

Magnetic doodle pad

Paper

Markers

Stickers

Scissors

Solar System Props

Styrofoam balls

Magnetic doodle pad

Paper

Markers

Stickers

Scissors

Appendix F

Post-Lesson Interview Protocol

Post-Lesson Interview Protocol

The post-lesson interview protocol outlined here was implemented two times, on two different days, with student participants from two classrooms.

Following a science investigation, each student participant was asked to tell the researcher a story about what they learned. This was referred to as either the "story interview" or the "interview." Following the first investigation, one half of the students in each class were interviewed prior to taking the teacher-driven assessment. The opposite half of the students were interviewed following the teacher-driven assessment. The student code numbers were assigned on day one, based on the order of their interview with the researcher. For example – the first student interviewed from the first class was assigned code SP1C1 (Student Participant 1, Classroom 1). The student participant codes (identification numbers) were permanent and did not change for the second investigation on day two. During the second investigation (day two), the order of the classrooms and the order of which assessment the students were asked to take first were both reversed. For example - Classroom 1 participated in the study first on day one, and the students from Classroom 1 who were interviewed prior to the teacher-driven assessment on day one, were interviewed after taking the teacher-driven assessment on day two.

Environment

The interview took place in the library area of the participating school. It included a table and four chairs in the middle of the library. The teacher-driven assessment was administered in the participating teacher classrooms. In the library, the researcher was seated across from, or next to, the student participant (their choice) while the remaining, interview first, students waited with the science support specialist in the science classroom which was located across the hall from the library. Resources on the researcher's table were – two audio recording devices, researcher's note pad, and materials that were offered as props to use during the storytelling process. These materials included – pencils, markers, blank paper, Matchbox cars, magnetic doodle pads, small rubber balls, and Styrofoam balls.

Narrative for the Introduction to the Student Participant

Researcher: *Hello, __________. How are you today? I would like to thank you for allowing me to interview you about what you learned today. I'm asking everyone who agreed to participate in my research to tell me the story of what they learned during the science lesson today. While you tell me the story of what you learned, I'm going to be taking notes and recording your voice on this tape recorder. If at any time you would like to use the story props you see on the table, to help tell your story, you may do so without asking. Okay? Do you remember the three parts of a story? Can you name them (student may need reminded of stories having a beginning, middle, and end)? Before we get started I wanted to ask if you have any questions for me (questions will be documented in the researcher's notes)? Are you ready to get started?*

The researcher turned the audio recording device to "on" following the narrative referenced above (note that the researcher used written anecdotes to capture demeanor, expressions, movement, etc.).

Steps Taken to Implement the Story Interview Session

1. Researcher asked – "Can you tell me a story about what you learned during the science lesson today?"
   i. May have asked the student to speak louder or into the microphone if needed.
   ii. May have prompted the student to continue with the story if he/she becomes distracted.
2. In conjunction with the audio recording device, researcher documented portions, and aspects, of the interview in writing.
3. Once the student was finished, the researcher repeated the story back to the student participant.
4. The researcher asked – "Is there anything you would like to add to your story now that I've read it back to you?"
5. If the student responded that he/she would like to add additional information, the researcher allowed time for that to transpire.
6. The researcher used phrases such as – "Are you finished with your story, or is there anything else you would like to add?"
7. Researcher's final statement – "Thank you for sharing what you learned today. To thank you for participating in my study I'd like to invite you to pick out a gift. Take any one thing from this bag (researcher pulled bag out from under desk).

Researcher guided each student participant back to the science classroom and waited for the next student participant to enter the library.

Appendix G

Teacher-Driven Assessment – Energy

ENERGY TEST (front page) **NAME__________________**

--------1. **What is energy?**

A. Energy is something that makes things jump
B. Light and sound are both energy
C. Energy is Matter that has substance
D. Energy is the ability to do work or cause change

-------- 2. **Which of these has kinetic energy?**

A. A ball on a shelf about to fall
B. A ball that has been thrown and is moving in the air
C. A ball on the floor not moving
D. A book that is sitting on a desk

-------- 3. **Which of these has potential energy?**

A. A basketball that is on a shelf about to fall
B. A football that has been thrown and is moving in the air
C. A spinning bike wheel
D. A person running outside

-------- 4. **Which of these is a source of mechanical energy?**

A. Lights
B. Wind
C. Food
D. Candy

-------- 5. **Chemical energy is stored in ______________?**

A. Moving objects
B. Wind
C. Food
D. Batteries

ENERGY TEST (back page) **NAME__________________**

-------- 6**. What is the main source of light energy on Earth?**

A. Candles
B. Street Lights
C. The Sun
D. House Lights

-------- 7. **Which of these has potential energy?**

A. A football flying through the air
B. A car rolling down a hill
C. A book sitting on the edge of the table about to fall
D. A basketball bouncing in the gym

-------- 8. **Rubbing your hands together creates what type of energy?**

A. Potential
B. Mechanical
C. Chemical
D. Heat

-------- 9. **Electromagnet energy is energy from__________________?**

A. Water
B. Electricity
C. Wind
D. Sun

-------- 10. **If an arrow is flying through the air, it has what type of energy?**

A. Kinetic
B. Mechanical
C. Nuclear
D. Potential

Appendix H

Teacher-Driven Assessment – Solar System

SOLAR SYSTEM TEST (front page)   **NAME**______________________________

--------1. **The Sun is a ____________________?**

A. Planet
B. Meteor
C. Star
D. Galaxy

-------- 2. **How does the Sun help the Earth?**

A. It gives us seasons
B. It cools the air
C. It keeps the Moon in place
D. It gives Earth heat and light

-------- 3. **Earth is a ________________________?**

A. Planet
B. Star
C. Galaxy
D. Meteor

-------- 4**. How many weeks does it take the moon to travel around the Earth?**

A. 1 Week
B. 2 Weeks
C. 3 Weeks
D. 4 Weeks

-------- 5. **An Axis is ________________________?**

A. Moving objects
B. An imaginary straight line that something turns around
C. A Planet
D. The name of a really good drummer

SOLAR SYSTEM TEST (back page) **NAME**________________________

-------- 6. **What is the main source of light energy on Earth?**

A. Candles
B. Street Lights
C. The Sun
D. House Lights

-------- 7. **The Earth travels around the ________________________?**

A. Moon
B. Sun
C. Planets
D. Solar System

-------- 8. **Astronauts have visited the __________________________?**

A. Sun
B. Mars
C. Moon
D. Jupiter

-------- 9. **The Moon changes as it goes through its _______________.**

A. Motions
B. Day
C. Monkeys
D. Phases

-------- 10. **How many Planets are in our solar system?**

A. 10
B. 8
C. 4
D. 1