12-1-2014

Student perspectives of classroom assessment

Kyle Hemje
University of Northern Colorado

UNIVERSITY OF NORTHERN COLORADO

Greeley, Colorado

The Graduate School

STUDENT PERSPECTIVES OF CLASSROOM ASSESSMENT

A Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree of
Doctor of Education

Kyle Hemje

College of Education and Behavioral Sciences
School of Teacher Education
Educational Studies

December 2014

This Dissertation by: Kyle Hemje

Entitled: *Student Perspectives of Classroom Assessment*

Has been approved as meeting the requirement for the Degree of Doctor of Education in College Education and Behavioral Sciences in School of Teacher Education, Program of Educational Studies.

Accepted by the Doctoral Committee

_____

Jennifer Harding-DeKam, Ed. D., Research Advisor

_____

Valerie Middleton, Ed. D., Committee Member

_____

Christine Moroye, Ph. D., Committee Member

_____

Steven Pulos, Ph. D., Faculty Representative

Date of Dissertation Defense _____

Accepted by the Graduate School

_____

Linda L. Black, Ed. D.
Dean of the Graduate School and International Admissions

ABSTRACT

Hemje, Kyle. *Student Perspectives of Classroom Assessment*. Published Doctor of Education dissertation, University of Northern Colorado, 2014.

This study used a convergent, parallel, mixed methods design to examine secondary math students' perspectives of classroom assessments. Converging both quantitative (student surveys) and qualitative (focus groups and interviews) data strands offered a more complete perspective of what classroom assessment strategy students preferred. Students heavily preferred a strategy named the "Dan Meyer" strategy for both tests and quizzes. Major qualitative themes emerged to support the disparity in views between both classroom assessment strategies found in this study. Since major differences existed between students on the classroom assessment strategies, further research is suggested to examine the views of secondary math students and which classroom assessment strategies help them learn.

ACKNOWLEDGEMENTS

First, I have to thank my family for inspiring me to finish this race: to Stephanie, my beautiful wife, for standing by my side and encouraging me through this stage of our lives. For over four years, you never let me be discouraged and you always reminded me to invest in things that really matter. Thank you for your love and support! I would not have finished without you standing next to me! I also have to thank my two boys--Titus and Marcus--both of whom were born during this process. Neither of you realize how inspirational you have been in helping me finish strong. I love you both!

I need to thank my entire committee for challenging and helping me through this process. I could not have been more blessed to have professionals give me their time whenever I needed it for guidance and counseling. Each of you has played a special role in helping me develop and become excellent in ways I did not think were possible.

Dr. Christy Moroye, thank you for challenging me and giving up your time to work with me. Your insights were clarifying and helpful in the final stages of my dissertation. Your willingness to work with me and challenge me to be excellent speaks volumes to your character.

Dr. Gary Fertig, thank you for serving me throughout my career. You have sacrificed more than most people could fathom just because you saw value in me and my future. Thank you for your grace and helping me (and Tina) take a class we both needed

and for not being paid for it. I hope to one day be as well-spoken and respected in higher education as I believe you are. Thank you for valuing me and my education.

Dr. Steven Pulos, thank you for your insights into both education and psychometrics. You have inspired me to strive to be an expert at the craft of education and understanding assessments. Thank you for sacrificing your time to help me get better. You are one of the best teachers I have ever met. Your brilliance and humor will not be forgotten.

Dr. Valerie Middleton, I owe you more than I can put into words. You have never stopped challenging me or trusting me in the five years I have known you and you genuinely have invested in me as a leader and peer. I can't express my gratitude enough. Working with you has brought me joy and helped me desire to work in higher education in a teacher preparation program. Thank you for your friendship and inspiration.

Dr. Jenni Harding-DeKam, you are one of the greatest people I have met in all of higher education. I am utterly amazed at how you have come along side me and taken me under your wing as a peer and a friend. Thank you for investing in me and believing in me when I did not. Thank you for being my chair and for always encouraging and challenging me every step of the way. Thank you for your insights, your humor, and for leading me in ways nobody else could have.

Last, I have to thank my parents and friends for all of their support throughout these five years of challenge. Mom and dad, you have been and continue to be wonderful models of what marriage looks like. Thank you for investing in what matters and loving Steph and me as we fought through these years to make it. Marc and Holly, thank you for loving us and giving up anything to help us make it. Tina Errthum, thank you for

encouraging me to finish my dissertation on time and for your awesome organization ideas. And to all my friends from my life group through my church, you have always supported and blessed me and my family. You have all been our family in Colorado. Steph and I love you very much!

DEDICATION

This dissertation is dedicated to my King and Savior Jesus Christ. You always have been and always will be my hope because your tomb is empty.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION TO THE STUDY

The No Child Left Behind (NCLB) Act of 2001 was passed by President George W. Bush in 2002. The signing of this law was the first time in over two decades the government determined the direction of education. Its purpose was to ensure that all students receive a high-quality education (NCLB, 2002, Section 1001). The goal of this bill was to identify schools that do not provide a high-quality education and transform them into superior learning communities. Accountability was measured through adequate yearly progress (AYP) and by assessments from each state toward academic achievement standards of the state (NCLB, 2002, Section 6161). To do this, the law placed accountability on the entirety of the educational structure, including not only educators but also administrators and district-level employees. The law required educators to be held accountable for student learning (NCLB, 2002, Section 3122) and the state would hold local educational agencies accountable for student achievement (NCLB, 2002, Section 6161) to make AYP. A key component of the NCLB legislation and associated emphasis on standards-based education was the use of standardized test scores to measure student achievement. By using standardized tests in all public school districts, accountability of all public teachers could be measured (NCLB, 2002, Section 6161). The NCLB statute resulted in schools and districts using test scores as data for accountability purposes. School districts, administrators, and teachers were then held

accountable for the learning gains of their students. Two content areas in particular, reading and math, were the subjects of intense attention by the law; now the testing scores in those two areas hold a nation accountable.

High-stakes tests are instruments used for accountability of educators by districts. These same data are used by the government to keep districts accountable. School districts understand the significant pressures of accountability and continually changing curricula to help enhance testing scores in response to those pressures. Since accountability testing can significantly affect the curriculum (Berliner, 2011), curricula and high-stakes testing in turn both have a direct impact on teacher assessment practices.

Assessments are now more important in educational settings and are used to measure different educational variables of interest (Popham, 1995). Popham (1995) explained that teachers use assessments to determine strengths and weaknesses of student comprehension, monitor growth of student learning, improve the assigning of grades, and help determine the effectiveness of instruction. These are all areas of necessary attention in classrooms for accountability purposes. Understanding what students know and need to know in order to be successful is vital to help teachers.

Educators use different types of assessments that include both summative and formative methodologies. Often times, these formative and summative assessments are designed specifically by teachers for classes. Each type of assessment has a different purpose and is valued in different ways by the teacher, administration, and government. Out of the three, teachers indicate classroom assessments are of greatest value to them because they are designed by the instructors and help guide future teaching in the classroom (Stiggins & Bridgeford, 1985). Of the three kinds of assessments used in

schools, classroom assessments should be researched most heavily as classroom assessments have the most influence on student learning (Allen, Ort, & Schmidt, 2009).

The forthcoming literature review illustrates general assessments, how formative and summative assessments are used in education, how classroom assessments are cumulative forms of assessment techniques, and classroom assessment strategies secondary mathematics teachers use. The literature review then examines both the teachers' and students' views of assessments in education.

Need for the Study

For teachers to know what is best for students and to help students be successful, assessment must take place on a consistent basis in the classroom (Stiggins, 2002). Assessment strategies that work best have been studied time and again, predominantly through the eyes of educators. Many highly disputed claims that teachers can use classroom assessment practices that raise test scores (Black & Wiliam, 1998b) to viewpoints of how timely assessments can be used to assess progress and attainment (Tiknaz & Sutton, 2006) have been done--all from the viewpoints of the teachers in the classroom. Teacher perspectives are important to obtain as the teachers have the craft, knowledge, and skills associated with successful design and implementation of classroom assessments. However, it is important to note the complete lack of literature concerning the students of these same classroom assessment strategies and techniques.

Student perspectives are still important as shown by the literature. However, almost all literature associated with the perspectives of students was generated by college level students. Many examples in higher education see student perspectives as beneficial and worth studying regarding effective teaching and assessment practices (Onwuegbuzie

et al., 2007: Seymore & Chance, 2010: Wren, Sparrow, Northcote, & Sharp, 2009; Yang, 2010); therefore, many of students in these studies only a few months prior were in secondary high schools. It could be argued that students' perspectives are just as important prior to going into an undergraduate program as they are within an undergraduate program. Indeed, looking at the literature, students who are in high school do not have much research backing any kind of perspective, especially one that considers looking at the classroom assessments they are so readily given on a daily or weekly basis. With such a push in educational society to give varied assessments and collect data on student learning, could it be possible that the classroom assessments teachers give so often may not be as beneficial as they think to helping students learn? Gaining insight from students as to what classroom assessments and strategies teachers use to benefit their learning would be a significant contribution to classroom teachers, administrators, and district officials. Although the benefits of as well as consequences from classroom assessments are greatest for students, virtually nothing is known about how students perceive the utility of such assessments.

Therefore, the purpose of this mixed methods study using Creswell and Plano-Clark's (2011) framework of a convergent parallel design was to gain secondary math students' perspectives of classroom assessments and to compare them with the perspectives of their teachers. A convergent, parallel, mixed methods design was used. Quantitative and qualitative data were collected in parallel, analyzed separately, and then merged using a joint-display comparison (Creswell & Plano-Clark, 2011, p. 223). Quantitative data in surveys given to students included frequency scales on classroom assessments and techniques used by their teachers. The qualitative data were derived

from focus groups with students and semi-structured interviews with teachers concerning assessment strategies not found on the survey and usefulness of practiced classroom assessment strategies. Collecting both quantitative and qualitative data allowed results to be merged for greater insight into the assessment strategies used in the classrooms and gave comparisons between students and teachers.

Research Questions

When reading previous literature, many quantitative and qualitative research questions emerged about assessments and how they were used. Because classroom assessments need to be explored further, a combination of quantitative and qualitative research questions were examined in this study. A mixed methods research question was utilized along with quantitative and qualitative research questions. According to Creswell and Plano-Clark (2011), "The combination of quantitative and qualitative data provide a more complete understanding of the research problem than either approach by itself" (p. 8).

Quantitative Research Question

Q1 Do differences exist between secondary students' perspectives of classroom assessments and their age, gender, or assessment strategy?

Qualitative Research Questions

Q2 What assessment practices are used most frequently in secondary math teachers' classrooms according to both students and teachers?

Q3 Which assessment practices do secondary math teachers consider to be most useful to their teaching practice and why?

Q4 Which assessment practices do secondary students consider most useful to their learning and why?

Mixed Methods Research Question

Q5 Are there differences between secondary math students' perspectives of classroom assessments and their age, gender, or assessment strategy, and how do students' perspectives compare with their secondary math teachers' perspectives of classroom assessment practices?

Definition of Terms

For the purpose of this study, the following terms are defined using definitions from literature.

Assessment. Defined by Nitko and Brookhart (2011), it is "a broad term defined as a process for obtaining information that is used for making decisions about students; curricula, programs, and schools; and educational policy" (p. 3).

Classroom assessments. "Teachers' assessment practices as distinct from any testing mandated by the district or state" (Young & Kim, 2010, p. 4).

Formative assessment. Popham (2006) defined the term as follows: "An assessment is formative to the extent that information from the assessment is used, during the instructional segment in which the assessment occurred, to adjust instruction with the intent of better meeting the needs of the students assessed" (Popham, 2006, pp. 3-4).

Summative assessment. The process by which teachers gather evidence in a planned and systematic way in order to draw inferences about their students' learning based on their professional judgment and to report at a particular time on their students' achievements (Harlen, 2005b, p. 247).

These definitions were used in this dissertation as they have were found to support one another and all came from reputable researchers who have helped shape the current assessment landscape in education.

Summary

This study addressed students' perspectives of the classroom assessments they were given in secondary math classrooms. Chapter II provides a framework toward understanding assessments in education and perspectives of both teachers and students through the use of current literature. Chapter III address the methodology and research design used in the study including a description and rationale for using both quantitative and qualitative methodologies. It also includes descriptions of the settings, participants, and instruments to be used.

CHAPTER II

REVIEW OF LITERATURE

A review of the relevant literature is presented in this chapter to gain a purposeful understanding of assessment and perspectives of assessment practices in classrooms. This study examines the perspectives of classroom assessments through the eyes of secondary mathematics students and their teachers. The first section of this chapter reviews relevant literature surrounding a definition of assessments and what they mean in education. It includes a discussion of two general types of assessment practices: formative and summative. The second section of this chapter includes a literature review about types of classroom assessments with a specific focus on mathematics assessments. The third section of this chapter reviews research associated with student and teacher perspectives of assessments. This literature review guided the researcher and justified the focus for this research.

Assessments

Defining assessment has proven to be difficult in education. Researchers have provided many definitions of assessment that attempt to explain what assessments are and the many ways they are used in schools. Popham (1995) suggested that assessment is "a formal attempt to determine students' status with respect to educational variables of interest" (p. 3). He also explained assessment as "a broader descriptor of the kinds of educational measuring that teachers do" and they "cover many more kinds of

measurement procedures" (p. 2). Taylor and Nolan (2005) suggested assessments are both tools and processes (p. 4). Allen et al. (2009) summarized the difficulties of settling on an assessment definition by saying, "Ultimately, assessment is teaching. As long as educators believe that teaching matters in the development of students' learning and lives, the development of teachers' assessment practices is too critical for anybody to go it alone" (p. 80). Allen et al. further suggested a framework to identify aspects of assessment including teachers' curriculum content, connection to learning goals, clarification of students' work criteria, providing support for students, encouraging development, and providing feedback. Using this perspective, it is easier to comprehend what qualifies as an assessment rather than how to define it. However, assessment can be applied to many aspects of education. For example, assessment can be a window into the curriculum (Holler, Gareis, Martin, Clouser, & Miller, 2008), meaning the style and use of assessments allow the opportunity to see the philosophies of the curriculum at work in the classroom. For the purposes of this research project, assessment followed the definition provided by Nitko and Brookhart (2011): "a broad term defined as a process for obtaining information that is used for making decisions about students; curricula, programs, and schools; and educational policy" (p. 3). It is significant to recognize assessments are complicated and they deal with many aspects of education.

Assessment is not only vital to education but is equal with teaching (Allen et al., 2009). Assessments are a major part of education. In fact, researchers suggested assessments go with education and education cannot be without assessment (Allen et al., 2009; Shavelson et al., 2008). Stiggins (2002) went so far as to say that assessment is evidence of success on behalf of students, teachers, and the educational system. Some

professionals in the education community argue that assessments provide important evidence for specific purposes that not only allow decisions to be made but permit understanding of how to collect data and interpret the data to occur (Harlen, 2005a). However, this understanding is not fully grasped in education. Heritage (2007) contended that although assessments are a part of education, many times the classroom teachers view assessments as competing with actual teaching. The argument is made based on the idea that teachers sometimes have specific assessments forced into their classrooms and these assessments take time away from teaching (Heritage, 2007). These assessments specialists identify assessments as a major piece of education today.

Students benefit from assessments in multiple ways including better communication, teacher-student involvement, and the potential for improved student learning. Assessments provide ways to communicate learning of students and provide a means to accomplish a task with step-by-step instructions (Allen et al., 2009). When teachers are also involved with feedback, monitoring, and constant assessment, entire classes have the potential to benefit (Heritage, 2010). Some researchers even argue assessments involving feedback have been shown to improve student learning (Black & Wiliam, 1998b). Although the amount of student learning gained has been contested (Dunn & Mulvenon, 2009; Kingston & Nash, 2011), there is still agreement that students benefit from assessment practices in the classroom.

Assessment is necessary in education for practical reasons such as data collection on students' learning. Different styles of assessments allow different ways of collecting data. In a major research study done by Stiggins and Bridgeford (1985) that studied assessments used in classrooms, their research concluded that different school subjects

required different data collection techniques and teachers should strive to collect data on student achievement. Reinforcing this inference were two later studies. The first study involved 262 National Council of Teachers of Mathematics (NCTM) members; 77.9% used traditional tests and 67.5% used quizzes as primary sources of data collection (Ohlsen, 2007). Compare this to a case study by Van Zoost (2011) who studied English classrooms and determined traditional assessment practices should be replaced by authentic or alternative performance assessments. All of these researchers identified that assessments provide necessary and meaningful data collection in education and different assessment techniques should be used in different subject areas.

In recent years, a push for data collection via assessments has come from a single all-encompassing assessment instrument--state tests. The No Child Left Behind Act of 2001 (2002) was the primary source to force the agenda into classrooms. Its purpose was to ensure all students receive a high-quality education (NCLB, 2002, Section 1001). The goal of the legislation was to identify schools that did not provide a high-quality education based on a standardized test "score" and transform them into superior organizations so scores would increase. Accountability through adequate yearly progress (AYP) via assessments was required. Each state had to show growth toward academic achievement state standards (NCLB, 2002, Section 6161). Stiggins (2002) pointed out that the government has faith in assessment as a tool for school improvement. However, he highlighted the reality that assumed standardized testing would have the same effect on every student is foolish. Regardless, the NCLB statute resulted in schools and districts using state test scores as data for accountability purposes.

An unfortunate theme rising from assessment in education was the lack of assessment skills by teachers. Researchers strongly agreed that teachers lack the proper training necessary for having quality assessments (Ayala et al., 2008: Guskey, 2003; Harlen, 2005b; Stiggins, 2002; Young & Kim, 2010). Allen et al. (2009) suggested one of the main reasons for this was due to teacher preparation programs teaching specifically to assessments in higher education settings and not public education settings. Other researchers suggested that teachers have preconceived notions of assessment practice for their classrooms and these notions have a negative influence on how assessments should actually be implemented in their classrooms (Ayala et al., 2008). Sometimes teachers simply misunderstand assessments (Black & Wiliam, 2004) while some researchers suggest this lack of assessment skills by teachers hurts students (Stiggins, 2002). The concerns brought forward by researchers emphasize the importance of better training for classroom teachers and how many teachers do not currently have the skills in place to improve assessments.

Research generally pointed out two different kinds of assessment practices in education--formative assessments and summative assessments. Both assessments can be used in different ways. Harlen (2005a) suggested formative and summative assessments be discussed separately because each has a different purpose: formative helps learning and teaching while summative records and reports. Summative assessments can also provide students and teachers with formative data uses (Allen et al., 2009) while formative assessments are primarily designed to give instant data about student knowledge so teachers can adjust their teaching to aid student understanding (Ayala et al., 2008; Britton, 2011). It is necessary to briefly conceptualize these kinds of

assessment practices in general terms in order to comprehend the different types of assessments used in classrooms, commonly known as “classroom assessments” in research.

Classroom assessments incorporate multiple kinds of assessment practices. Young and Kim (2010) pointed out the fact that teachers do not rely on one source of information when using assessments and many forms of formal and informal styles are used. Classroom assessments are predominantly teacher-developed and used by teachers to evaluate their students (Boothroyd, McMorris, & Pruzek, 1992). Overall, formative and summative assessment practices are used in a multitude of ways in classroom settings.

Assessments are generally difficult to define as a whole but the literature supports assessments that not only belong in education but are a vital part of it. Assessments can benefit both students and teachers--students by guiding their learning and teachers by using data to guide instruction. There is still a dominant push toward high-stakes testing; concerns from professionals in education have brought up possible deficiencies of current assessment practice done by teachers (Black & Wiliam, 1998b, 2004; Boothroyd et al., 1992: Cizek, Rachor, & Fitzgerald, 1995; Harlen, 2005b; Young & Kim, 2010).

Formative and summative assessments are two general kinds used by teachers in classrooms and most assessment practices fall into one of these two categories. This discussion begins with formative assessments followed by summative assessments.

Formative Assessment

The term “formative” was originally given to assessment practices by Michael Scriven (1979) and is widely agreed upon by the educational field (Dunn & Mulvenon,

2009; Newton, 2007; Tarus, 2005). In the 1960s, Scriven used the term "formative" when he was evaluating the roles and goals of assessment use in education. Popham (2006) credited Scriven with the "earliest and most influential evaluation-specific constructs to be accepted by America's emerging collection of educational evaluators" (p. 2). Popham pointed out the significance of not only deciphering two roles for evaluating but also for identifying the purpose of the formative evaluation--still having time to adjust and improve. According to Tarus (2005), Scriven argued that assessment is a single process no matter what is being assessed. This process also included a judgment based on a standard (Tarus, 2009). Thus, Scriven is widely agreed as the researcher who identified different assessment practices with different purposes.

However, agreement on the definition of formative assessment in education is continuously being debated amongst scholars. This lack of agreement on the terminology (Taras, 2005) has led to numerous interpretations. Formative assessment has largely been defined by some scholars as a practice to adjust instruction (Black & Wiliam, 2004; Dunn & Mulvenon, 2009; Fluckiger, Vigil, Pasco, & Danielson, 2010; Heritage, 2007, 2010; Hwang & Chang, 2011; McTighe & O'Connor, 2005). Many researchers point out how formative assessments provide feedback to students and teachers and help students improve (Allen et al., 2009: Dunn & Mulvenon, 2009; Fluckiger et al., 2010; Heritage, 2007; Hwang & Chang, 2011: McTighe & O'Connor, 2005: Tiknaz & Sutton, 2006). Some researchers identify formative assessments as being "snap shots" of student learning (Ayala et al., 2008) while others feel formative assessments promote learning in the classroom (Black & Wiliam, 1998a; Chappuis & Stiggins, 2002). Dunn and Mulvenon (2009) made the case that the use of the results determined if the assessment

practice was in fact a formative one. With all the minute disagreements, some leading experts in the field of formative assessment such as the State Collaborative on Assessment and Student Standards (SCASS) and the Formative Assessment for Students and Teachers (FAST) came together to define formative assessment as "an assessment is formative to the extent that information from the assessment is used, during the instructional segment in which the assessment occurred, to adjust instruction with the intent of better meeting the needs of the students assessed" (Popham, 2006, pp. 3-4). The definition uses "instructional segment," which means a short period of time--probably a class period or lecture. For the purposes of this research study, this definition of formative assessment was used.

Formative assessments are primarily made by teachers (Boothroyd et al., 1992) and frequently used in classrooms (Boothroyd et al., 1994: Young & Kim, 2010). In a study done by Cizek et al. (1995), a strong trend showed 84.6% of teachers developed their own minor, or formative, assessments. Ohlsen (2007) also conducted a study showing teachers preferred to create their own formative assessments for use in their classrooms. Young and Kim (2010) also pointed out that teachers favor spontaneous assessments, i.e., teachers prefer to provide additional ways to represent concepts when teaching students based on formative evaluation during instruction. Formative assessment strategies are not only designed by teachers but are typically done in the classroom and are central to teachers' instructional practices (Allen et al., 2009). Chappuis and Stiggins (2002) went so far as to support the idea that formative assessments not only happened often in classrooms but they should be happening

continuously. In summary, formative assessments are favored by teachers and are used often in the classroom so there must be reasons for having them present.

One such reason would be because formative assessments benefit teachers. Fluckiger et al. (2010) claimed formative assessments help teachers improve their teaching. Ginsburg (2009) explained it as helping the teacher plan effective instruction. He believed formative assessments help the teacher plan not only for a class but for individual students. He also made the point that formative assessments could provide insight as to how a student thinks. According to Heritage (2007), formative assessments provide information to teachers to move learning forward and Stiggins (2002) made the case that formative assessments provide teachers with data to make instructional decisions. Formative assessments guide teachers to improve the learning of their students (McTighe & O'Connor, 2005). Ultimately, formative assessments can provide many benefits for educators.

Similarly, students also have the potential to benefit from the use of formative assessments in the classroom. Formative assessments provide students with feedback (Allen et al., 2009: McTighe & O'Connor, 2005), focus students on progress being made (Chappuis & Stiggins, 2002; Harlen, 2005a), and help guide students to what is necessary to reach their learning goals (Chappuis & Stiggins, 2002; Fluckiger et al., 2010: Heritage, 2007). Chappuis and Stiggins (2002) proclaimed involving students in formative assessment practices could encourage achievement, which could lead to increased confidence and learning. Perhaps more simply put, appropriate formative assessments help students learn more (Stiggins, 2002).

The literature showed that formative assessments are important to both teachers and students. The data collected are for both the teacher and the student (Chappuis & Stiggins, 2002; Heritage, 2007). Formative assessments involve gathering data from students on their progress (Britton, 2011; Heritage, 2007) and using the data in a proper way to better support student learning (Harlen, 2005a). Formative assessments provide data showing what students know and can do at a specific point in time (Ayala et al., 2008); this interpretation of the data provided is recognized by most educators. Because of the dynamics of the data formative assessments provide teachers, educators are able to track the understanding of their students in a unique way (Ayala et al., 2008) that larger summative assessments cannot. Formative assessment data can and should be used by more than just educators as students can also benefit from the insights provided on their learning.

Much of the data provided by formative assessments have been credited for increased communication not only between teachers and students but parents as well. Providing feedback is one of the primary ways communication happens. According to Fluckiger et al. (2010), “Feedback is a key strategy of formative assessment” and that “feedback must be specific, simple, descriptive, and focused on a task” (p. 137). Black and Wiliam (1998a) pointed out that the feedback has to be useful. Popham (1995) claimed that “the use of formative assessments helps clarify the teachers’ instructional intentions” (p. 10) to the students. Clarification provides feedback at multiple levels for the students and provides collaboration with teachers to develop a shared understanding of student learning (Heritage, 2007). Black and Wiliam (1998b) indicated the interactions between teachers and students should be thoughtful, reflective, focused, and

performed so all students should be capable of sharing and expressing ideas. Just as feedback on formative assessments provides opportunities to communicate between teachers and students, formative assessments help communication with parents. Establishing an obvious and complete set of achievement expectations with quality assessments can lead to stronger communication with parents (Stiggins, 2002). Taylor and Nolan (2005) said parents and guardians will need to understand both teacher goals for their children and the child's progress toward those goals. Formative assessments can help communicate learning and improvement in learning to parents, teachers, and students.

As the literature on formative assessments grows, so do some of the claims and controversies regarding the impact formative assessments have on education. Recently, there have been strong claims that formative assessments, when designed and implemented correctly in education, can significantly increase student academic achievement. In a meta-analysis conducted by Black and Wiliam (1998b), they claimed to have identified formative assessments; the feedback they provided was able to yield substantial learning gains as high as .4 to .7 effect sizes. These were significantly higher than most educational interventions (Black & Wiliam, 1998b). Perhaps more impressive were the claims by Black and Wiliam (1998b) that formative assessments helped lower achievers more than other students. Later, Black and Wiliam (2004) went so far as to say their research was "strong and rigorous" and it "provides evidence that improving formative assessment can raise standards of students' performance" (p. 20). The evidence produced by Black and Wiliam has been sited over 2,000 times (Kingston &

Nash, 2011) by other professionals in the field. With such significant claims being made in favor of formative assessment, backlash to such strong claims have begun to emerge.

As indicated from the literature that formative assessments can increase student achievement, growing research has argued against the conclusions of Black and Wiliam (1998a). Dunn and Mulvenon (2009) convincingly pointed out a myriad of research used by Black and Wiliam (1998a) in which conclusions were inappropriately interpreted. Such examples included generalizing findings of studies that were done predominantly with students identified with special needs, inappropriate interpretation of a study that looked at increased numbers of assessments, and using results of formative assessment practices designed specifically for impoverished kindergarten students and applied to the general population of students (Dunn & Mulvenon, 2009). Kingston and Nash (2011) also pointed out methodology flaws performed in the research by Black and Wiliam (1998a) including the fact that the meta-analysis they performed did not use any quantitative meta-analytic techniques on their data collection. The effect sizes of .4 to .7 Black and Wiliam (1998a) concluded were possible for learning gains might in fact be unfounded (Kingston & Nash, 2011). Finally, it is also important to note that the original authors later indicated the quantitative results of their study were, in fact, difficult to interpret (Wiliam, Lee, Harrison, & Black, 2004), which is unusual for authors to reach such a strong conclusion and later contradict their findings. With the aforementioned studies being critical of the original argument that formative assessments help students learn, the debate of how formative assessments affect students is still ongoing.

Some studies still support formative assessments on the premise they still might improve student learning. Heritage (2010) made the point that teachers need to lead the

learning done in their classrooms and not simply react to it. Heritage argued the best way to do this is to use formative assessments as guides for both teachers and their students to work within the zone of proximal development. Although Kingston and Nash (2011) criticized the work of Black and Wiliam (1998a), they also pointed out through their own meta-analytic approach of all work done on formative assessments after 1988, the effect sizes were much smaller than original estimates by Black and Wiliam (1998a)--.20 being the weighted mean effect size and .25 being the median of the observed effect sizes. However, they did remark, "Results, though, do indicate formative assessment can be a significant and readily achievable source of improved student learning" (Kingston & Nash, 2011, p. 33). With this most recent literature in mind, there is still evidence that formative assessments can still be beneficial to students in improving learning.

Overall, formative assessments are still relatively new in public education over the course of history. However, state officials and district administrators are quickly learning their importance in the era of accountability and that significance is beginning to trickle down to educators. Teachers now realize the benefits of formative assessments for themselves and their students and acknowledge formative assessments have the potential to not only aid in instruction but aid in student learning.

Summative Assessments

The definition of summative assessment is generally agreed upon in the literature and is noted as having a significantly different purpose than that of formative assessments. First, summative assessment is considered a process (Taras, 2005) in which students summarize what they have learned (Heritage, 2007). Tiknaz and Sutton (2006) pointed out summative assessment only records students' accomplishments at a particular

time. An interesting perspective offered by Dunn and Mulvenon (2009) suggested "the actual methodology, data analysis, and use of the results that determine whether an assessment is formative or summative" (p. 2). Similarly, summative assessments summarize what students have learned after a segment of teaching occurs (Linn & Miller, 2005: McTighe & O'Connor, 2005; Nitko & Brookhart, 2011). However, the definition provided by Harlen (2005b) seemed to summarize what much of the research conveyed:

> The process by which teachers gather evidence in a planned and systematic way in order to draw inferences about their students' learning, based on their professional judgment, and to report at a particular time on their students' achievements. (p. 247)

This definition of summative assessment was used for the duration of this research project. Summative assessments have a very different purpose than that of formative assessments and their purpose is generally agreed upon by those in the literature.

Interestingly, the term "summative assessment" can wrongfully be associated strictly with high-stakes assessments such as state tests given annually. The high-stakes summative assessments given in each state get most of the attention and resources in schools (Black & Wiliam, 1998b; Frey & Schmitt, 2010; Stiggins & Bridgeford, 1985). The amount of importance placed on state tests can mitigate the influence of other assessment practices (Wiliam et al., 2004). Associating high-stakes tests as a key summative assessment can have additional flaws such as not providing teachers with instructional information they can use (Harlen, 2005a; Heritage, 2007) to produce desired outcomes such as problem-solving and critical thinking (Harlen, 2005a). Overall, when wrongly associating the term "summative assessment" as a high-stakes test, the emphasis of these large tests is on student learning and achievement as a single event (Ohlsen, 2007) and not over the course of time.

Although many associate high-stakes tests with “summative assessment,” the reality is most summative assessments are made by teachers. In a study done by Cizek et al. (1995), which surveyed teachers regarding their assessment practices, 71% of teachers preferred to create their own summative assessments. Teachers’ pedagogical knowledge makes assessments more manageable and sustainable (Tiknaz & Sutton, 2006). Teachers who create their own summative assessments provide increased internal and external validity of the scores delivered by summative assessments (Harlen, 2005a), meaning the scores from the summative assessments are more trustworthy and useable. Teachers primarily use their own summative assessments for logical purposes--to capture scores for the purpose of grading their students (McTighe & O’Connor, 2005). Summative assessments are assessments designed and used by teachers with the specific purpose of assigning grades to their students.

As previously mentioned, the relationships between formative and summative assessments should align and summative assessments can guide students to understand ways to be successful on all assessment practices in the classroom (Shavelson et al., 2008). Summative assessments also help to index student achievement against standards when aligned (Shavelson et al., 2008). Summative assessments have the ability to provide both students and teachers with data on how to improve student learning (Allen et al., 2009) when they are aligned with standards, meaning summative assessments reveal what standards students have met. When students have an opportunity to “rehearse their understanding of new concepts” on the formative assessments before taking the summative assessment, students are more “aware of the content on which they will be assessed” (Britton, 2011, p. 17). Summative assessments are a critical and

cumulative piece of associating formative and summative assessment strategies in classrooms.

Summative assessments should be the principal ways educators measure what students have learned over a segment of time by aligning them with standards. Summative assessments are the primary ways students are measured to state standards (Tiknaz & Sutton, 2006) and achieving learning standards is at the front of sound educational practices (Lalley & Gentile, 2009). Guskey (2003) simply stated summative assessments should line up with state or district standards while Newton (2007) and Taras (2005) both pointed out summative assessments help derive standards-referenced judgments that are appropriate when assigning grades to students. Summative assessments help monitor progress toward students meeting state standards (Shepard, 2005) and they "provide evidence of success on the part of students, teachers, and the system" (Stiggins, 2002, p. 759). By using standards, summative assessments are the primary ways that educators measure students for academic learning and judgment.

Similar to formative assessments, summative assessments should be a primary source of data collection for teachers as they can help with communication between educators and learners. Summative assessments can provide data that have many purposes (Dunn & Mulvenon, 2009; Harlen, 2005a; Taras, 2005) in the classroom. Summative assessments provide teachers and students with data about learning (Allen et al., 2009; Dunn & Mulvenon, 2009; McTighe & O'Connor, 2005) and those data are used as feedback to help students learn (Black & Wiliam, 1998b). However, data from summative assessments are not always used properly. For example, summative assessments are sometimes used for the sole purpose of ranking students and not

collecting data (Heritage, 2007). Harlen (2005a) claimed summative assessment data could negatively motivate students due to low achievement. Nevertheless, summative assessments provide data and those data are valuable to teachers; much of the time, the data help communicate learning achievements to students.

Summative assessments collect data for teachers and students; the data can be collected in a multitude of ways because summative assessments do not have to be paper-and-pencil assessments. In an article almost two decades ago, Robert Marzano (1994) argued for performance-based summative assessments. He contended performance-based summative assessments provide data about students' ability to think, unlike traditional summative assessments. Marzano also suggested they "provide data on skills and abilities on various proficiencies" (p. 46). He provided evidence that 67% of the teachers who conducted the survey identified that performance assessments were helpful and they could provide better assessment information than paper-and-pencil tests. More recently, Britton (2011) pointed out alternative summative assessments, meaning summative assessments that do not require the use of pencils and paper, are means to demonstrate mastery of skills and abilities not normally assessed by paper and pencil tests (p. 16). These same alternative-style summative assessments could demonstrate content knowledge as well as show mastery of skills and concepts (Britton, 2011). The data these alternative types of summative assessments provide teachers could be valuable and help determine what students have learned in untraditional ways.

Summative assessments have much more agreed upon definitions than their counterparts. Summative assessments many times are high-stakes associated but the reality of education reveals that most summative assessments are, in fact, teacher-made.

They are designed to measure what students learn over a set amount of time to provide data to teachers on their students.

Finally, a relationship between summative assessments and formative assessments should exist. Harlen (2005a) indicated that formative and summative assessments have different purposes but similar processes. When these processes are aligned, everyone involved with the assessments benefits. Shavelson et al. (2008) recognized, "When the formative and summative function of assessment are aligned so that the signals about what counts as achievement are consistent to educators, students, parents, and the public, assessment is expected to improve student learning" (p. 297).

When both kinds of assessments are aligned, the results could potentially enhance student learning and provide opportunities for students to know what is required of them on all assessment practices including standardized tests (Shavelson et al., 2008). However, some researchers warned that formative and summative assessments should be similar; if not careful, they cannot be distinguished from one another (Young & Kim, 2010). A relationship between formative and summative assessments should exist; unfortunately, the balance between them is delicate and can be confusing if both are not fully understood.

Classroom Assessments

"Understanding the dynamics of classroom assessment is essential for improving education or even for understanding the current state of learning in classrooms" (Brookhart & Durkin, 2003, p. 28). Classroom assessments can only be understood once formative and summative strategies are grasped. Classroom assessments are not necessarily different from formative and summative assessments but must be defined to

better understand their intentions in the classroom. Indeed, classroom assessments get less attention than high-stakes tests (Stiggins & Bridgeford, 1985) and are often argued to be more important. Classroom assessments are made up of both formative and summative assessments and have the same intention--to inform and guide teachers in instructional techniques (Frey & Schmitt, 2010; Ohlsen, 2007). The difference is that many classroom assessments can be practiced daily (Brookhart & Durkin, 2003). Ohlsen (2007) pointed out many purposes to classroom assessments including grading, identification of student special needs, motivation, and monitoring instructional effectiveness (p. 5), which could all be daily occurrences. However, even though some classroom assessments happen during instruction with a specific purpose, they do not have to affect student grades (Frey & Schmitt, 2010). The different styles and techniques used by teachers can be quite vast (Brookhart, 1997; Ohlsen, 2007). For the purposes of this study, classroom assessments are readily defined by Young and Kim (2010) as "teachers' assessment practices as distinct from any testing mandated by the district or state" (p. 4). Using this definition provided a guide to understanding the concepts and components of classroom assessment necessary in education.

Classroom assessments use both formative and summative strategies of assessments. In a literature review done by Young and Kim (2010), they revealed classroom assessments combine formative and summative assessments to serve many functions. They pointed out multiple forms of each assessment style are performed by educators because teachers understand one assessment is not adequate enough as a source of information of student performances. Assessments of all kinds are done frequently and there is evidence to support teachers prefer to design them. Lastly, Young and Kim

concluded “assessment practices reflect their [teachers’] understanding of students’ learning processes and the content they teach” (p. 9). Brookhart and DeVoge (1999) also supported the idea classroom assessments can vary both formatively and summatively. Ohlsen (2007) found teachers reported using a variety of different classroom assessment strategies while they were teaching, suggesting both formative and summative assessment strategies were used. Brookhart (1997) concluded the more variety used for classroom assessments, the more meaningful those assessments may be for students. Classroom assessments are, at their core, summative and formative strategies.

Because classroom assessments are made up of formative and summative strategies, they are the most widely used assessments in education (Brookhart & Durkin, 2003). Black and Wiliam (1998b) pointed out both styles of classroom assessments are done in classes where learning takes place. In a study by Boothroyd et al. (1992), the researchers analyzed teachers’ measurement training and knowledge and concluded teachers gave classroom assessments frequently. For example, the math and science teachers in the study gave summative style assessments once every two weeks on average. Classroom assessments were predominantly done in a paper-and-pencil tradition (Frey & Schmitt, 2010) while summative tests and formative quizzes were the strategies of choice (Ohlsen, 2007) by most educators.

Classroom assessments are also the most meaningful assessments used by teachers. Classroom assessments can measure, report, and promote learning (McTighe & O’Connor, 2005) when teachers choose not only when and how to assess their students but also when to use information gathered from assessments (Young & Kim, 2010). Because teachers choose the assessment techniques they want to use, they tend to trust

their assessment choices because of the relationship the assessments have with instruction (Guskey, 2003). Additionally, Guskey (2003) claimed classroom assessments align with instruction and ideally should align with state and district standards. In addition, well-designed classroom assessments help guide students on their learning by giving them a sense of what they know and where they need to improve (Brookhart & Durkin, 2003; McTighe & O'Connor, 2005). Further evidence that classroom assessments are more meaningful to educators lies in the study by Boothroyd et al. (1992) who concluded teachers put more weight on their own assessments for grading students than any other factor. The literature identified classroom assessments as the most useful and purposeful assessments teachers could use.

The literature also identified classroom assessments as being able to provide feedback on student learning. Guskey (2003) argued, "The assessments best suited to guide improvements in student learning are the quizzes, tests, writing assignments, and other assessments that teachers administer on a regular basis in their classrooms" (p. 7). Classroom assessments have been identified as being able to guide student learning (McTighe & O'Connor, 2005); using classroom assessments to give feedback might help increase a desire to learn in students (Chappuis & Stiggins, 2002). Additionally, Brookhart & DeVoge (1999) argued for classroom assessments saying, "The way teachers communicate their expectations to students and the way they provide feedback as to how well these expectations are met help students form concepts of what is important to learn and how good they are learning" (pp. 409-410). Yet, some studies warn of teacher feedback not promoting learning at all. For example, Cizek et al. (1995) warned teachers' comments can at times focus on non-cognitive outcomes when

providing feedback. Still, classroom assessments that are properly designed and give proper feedback positively affect student learning (Frey & Schmitt, 2010).

With classroom assessments slowly being identified as important in the era of high-stakes testing, growing literature has pointed toward shortcomings of educators in assessment training. Heritage (2007) said, "Teachers learn how to teach without learning much about how to assess" (p. 141). Cizek et al. (1995) advised that teachers lack the ability to determine the quality and dependability of information given by classroom assessments. Traditional classroom assessments can be out of balance (Frey & Schmitt, 2010) and teachers may not see assessments as important to their instruction due to lack of training (Young & Kim, 2010). Guskey (2003) even suggested that lack of training could have damaging effects:

> When no suitable assessments are available, teachers construct their own in a haphazard fashion, with questions and essay prompts similar to the ones that their teachers used. They treat assessments as evaluation devices to administer when instructional activities are completed and to use primarily for assigning students' grades. (p. 7)

Ayala et al. (2008) remarked that teachers need professional development to better understand the value of classroom assessments with particular focus on formative assessment strategies. The benefits of appropriate training could lead to improved classroom assessment accuracy and efficiency (Chappuis & Stiggins, 2002). According to Boothroyd et al. (1992), there is an overall lack of teacher competency regarding measurement knowledge and values:

> Teachers are not routinely instructed in ways to collect and interpret information. …most teachers have not been adequately trained in how to develop and interpret a classroom test, even though these tests are the primary basis for assigning course grades and a major basis for a plethora of educational outcomes. (p. 8)

Research does point out some of the major concerns with teachers' lack of training; unfortunately, this leads to other assessment trepidations.

Flowing from the lack of teacher training in assessment comprehension are worries of reliability and validity of classroom assessments in the research. Unreliability and biases exist in teachers' assessments (Harlen, 2005b). Supporting a lack of reliability in teachers' classroom assessments was a study done by Cizek et al. (1995) in which the extent of measurement training obtained by science and math teachers was studied. It revealed that "assessment practices varied widely and unpredictably with no apparent relationship to characteristics such as practice level, years of experience, gender, or familiarity with district policies" (Cizek et al., 1995, p. 20). In other words, the assessments proved to be quite unreliable. A study by Boothroyd et al. (1992) examined the extent of measurement training obtained by 41 seventh and eighth grade science and math teachers and found only 20 had taken measurement courses. Of those 20, only three had ever critiqued classroom test items or constructed tests for classroom use, thus concluding the assessments being used by the educators in this study were unreliable. A further concern lies with lack of validity of classroom assessments used by teachers. Young and Kim (2010) revealed, "Assessment specialists are concerned about the validity of teacher-made assessments" (p. 27). This is rather alarming as the decisions made based on the classroom assessments students are given are rather "astronomical" and "validity is a concern" (McMorris & Boothroyd, 1993, p. 336). Many teachers do not understand how to properly interpret the scores of the assessments they give because "teachers' knowledge is not sufficient" (Boothroyd et al., 1992, p. 7). Unfortunately, teachers have a narrow view of their classroom assessments (Harlen, 2005b) since they

do not have measurement techniques to create proper classroom assessments (McMorris & Boothroyd, 1993). Therefore, the reliability and the validity concerns brought about by experts in the field are convincing.

Yet, classroom assessments are still argued to be best for students (McTighe & O'Connor, 2005). "Classroom assessment is critical to how a student is learning and how to best support that student's academic performance" (Allen et al., 2009, pp. 72-73). Teachers who assess for learning do so by means of classroom assessments, which can help with responsibility, engagement, and direction for students (Brookhart & Durkin, 2003; Chappuis & Stiggins, 2002). Brookhart and DeVoge (1999) said, "Classroom achievement is defined as accomplishing the instructional objectives intended in the teacher's plans. This achievement is conventionally measured with classroom assessments teachers construct or select for this purpose" (Brookhart & DeVoge, 1999, p. 411). When classroom assessments become a part of instruction, students will benefit (Guskey, 2003).

Evidence similarly supports that classroom assessments are advantageous for teachers as well (Guskey, 2003). A study done by McMorris and Boothroyd (1993) revealed that 69% of the teacher respondents identified the primary reason they gave tests was to assess student mastery of the content taught in class. Just as teachers use assessments to help understand students, assessments also help educators gather data on their students. Classroom assessments help teachers gather important data and assist them in deciding how to use their data (Young & Kim, 2010). Another reason classroom assessments help educators is because they guide instruction. McTighe and O'Connor (2005) indicated classroom assessments direct teacher instruction. Similarly, Guskey

(2003) argued classroom assessments help teachers identify what they taught well and where they can improve. Frey and Schmitt (2010) took instructional change one step further when they claimed, “Teachers can utilize the information they collect from their assessment of student learning to make adjustments in instruction” (p. 108). Ohlsen (2007) continued, “Assessment should be used to inform and guide teachers in their future instructional strategies” (p. 6). Teachers use their classroom assessments to gather data on students and to help their own instructional pedagogy.

Classroom assessments provide ways teachers can express what is important in their classrooms for their students. Teachers achieve this by choosing the focus of classroom assessments. Brookhart and DeVoge (1999) argued that teachers communicate their expectations on what students are supposed to learn via their classroom assessments. Their communications can be explicit or implicit and establish purpose for the assessment (Brookhart & DeVoge, 1999). Additionally, educators communicate mindsets in the classroom assessments they use. By choosing certain classroom assessments, teachers connect attitudes toward the subject matter students are to learn (Brookhart, 1997). Conversely, classroom assessments can also send the wrong message to students. Alkharusi (2008) conducted a study that examined over 1,600 ninth grade students and 83 science teachers in Oman to assess the effects of classroom assessment practices on students’ achievement goals. Alkharusi concluded there could also be a detrimental effect on students if classroom assessments emphasized public grades more than individual improvement. Educators must be careful in their classroom assessment choices as the assessments convey meaning of importance to students in positive and negative ways.

Finally, the literature revealed mixed results when looking at empirical results of teacher-made summative tests. Stiggins and Bridgeford (1985) completed a study exploring teacher-made assessments by surveying 228 teachers who taught in grades 2, 5, 8, and 11. They found about half of the teachers self-reported the use of teacher-made tests; this was constant across subject areas. The teachers who did not use published tests found them to be time-consuming and not meeting instruction or instructional needs. These researchers reported the higher the grade level, the more teachers reported using their own major tests (Stiggins & Bridgeford, 1985). Cizek et al. (1995) looked at a sample of 143 Midwestern elementary and secondary school teachers who self-reported assessment practices and found 74.2% developed their own major tests and 84.6% developed their own major quizzes. The researchers were surprised to find that 39% reported using commercial tests; they were also stunned at the data that reported newer teachers with less than five years of experience developed their own tests and quizzes more often than did veteran teachers. These authors reported that teachers lacked professional collaboration on assessment matters and concluded assessment practices varied wildly and unpredictably (Cizek et al., 1995). Another study conducted by Frey and Schmitt (2010) analyzed 140 teachers in grades 3 through 12 and had them self-report percentage of time used on various assessments in their classes. These authors looked at gender and grade levels as independent variables and the percentage of the time teachers used tests they made themselves as the dependent variable. Both grade level and gender had main effects: $F(2,113) = 4.21$, $p = .02$ and $F(1,113) = 4.76$, $p = .03$, respectively. They also found the interactive term was not significant and concluded male teachers used tests they created more often than did female teachers: $M = 67.53\%$,

SD = 32.29 for male educators and *M* = 50.93%, *SD* = 32.18 for female educators. Similarly, Ohlsen (2007) reported 262 high school National Council of Teachers of Mathematics (NCTM) teachers self-reported the frequency of the student assessment methods they used in the classroom to determine student grades and concluded that "teachers in this study showed a strong preference for creating their own classroom assessments rather than relying on publisher created student assessments found in instructional manuals" (p. 10). Studies revealed mixed results of major classroom tests and how often they were used for grading purposes.

Mathematics Assessments

Literature showed a need to further study mathematics assessments due to the necessity of their use by teachers, the number of variety and techniques used, and the aspects of working with primarily traditional practices. There were also deficits in mathematics assessments including teacher concepts of how to use assessments and how subjective math assessments really are in education.

Adams and Hsu (1998) gave a compelling argument as to why mathematics assessment is so important to study by commenting, "This focus on assessment is important to the development of mathematics instruction and curricula supporting children's learning" (p. 174). They continued expressing that teachers gain information to improve student learning and assessment is a primary factor to determine what and how students learn. Bahr (2007) stated math educators "need to assess a wider range of mathematical abilities" (p. 33). Teachers need an understanding of the cognitive levels of their students (Berenson & Carter, 1995) and most classroom teachers face a great diversity of mathematic ability levels (Bahr, 2007). Math educators need to carefully

consider what they are assessing and if it is a true indication of mathematical knowledge (Greenlees, 2011). Maccini and Gagnon (2006) also suggested it is important to identify instructional practices and assessment accommodations to help students be successful. Math educators need assessments in their classrooms for a multitude of reasons specifically related to understanding their students and helping them improve.

Mathematics teachers use a variety of techniques to assess their students. Math cannot be restricted to assessing students through conventional paper-and-pencil means (Clarke, 1992). Clarke (1992) argued, "It is also becoming accepted that the effective monitoring of students' learning requires the use of several different modes of assessment" (p. 29). Good alternative assessments can indicate to teachers how students think and comprehend (Berenson & Carter, 1995). Educators today must consider a large array of assessment techniques in their classrooms (Adams & Hsu, 1998). In the study by Adams and Hsu (1998), 269 math teachers were asked to identify a variety of assessment techniques used in their classrooms. The study showed teacher observations were the highest rated with the smallest standard deviation ($M = 4.753$, $SD = 0.488$) on a 5-point Likert scale. The teachers in the study also considered 20 different assessment techniques as being valid ways to assess students in math classrooms. In the same study, teachers also identified the use of essays in math classrooms as an assessment strategy that was lacking in frequency but was still used nonetheless (Adams & Hsu, 1998). Even interviewing was considered a way to further understand math students' processing (McIntosh, 1997). Math teachers used multiple strategies for assessing their students.

Traditional assessments in math classrooms were debated in the literature. Even though literature told of large numbers of assessment techniques backed up with

professional development to support teachers in these alternative assessments, the use of assessment strategies in math classrooms was still narrow (Warren & Nisbet, 2001). Berenson and Carter (1995) even argued a good traditional test is preferred to a poorly made alternative assessment. Yet these same authors contended, "Students' conceptual understandings cannot always be measured by traditional, short answers and paper-and-pencil tests" (Berenson & Carter, 1995, p. 184) and that the use of traditional assessments could lead to students pursuing grades rather than learning. Yet teachers continue to practice traditional means of assessment (Schmidt & Brosnan, 1996). Traditional assessments will most likely remain as the dominant assessment strategy in math classrooms, although their importance is still debated.

Assessments in mathematics classrooms are not without their share of deficits. Adams and Hsu (1998) pointed out that teachers' concepts may not be aligned with mathematical leaders in education. To further support this, Brown and Hirschfeld (2007) reported conflicting concepts of assessment were evident. Alternatively, assessments in math can be delivered at an incorrect ability level--being too difficult or too easy (Bahr, 2007) while Watson (2006) claimed subject-specific gaps exist in assessment strategies. To complicate matters further, Romagnano (2001) argued, "All assessments of students' mathematical understanding are subjective" (p. 31); a simple quiz can provide detailed information about what students know or little information at all depending on how it is designed, scored, and used. Even with these arguments highlighting the areas of weakness in math assessment, further research is needed to address these issues while strengthening the claims of mathematical assessment importance.

Perspectives on Assessments

This section of the literature acts as a review of the perspectives of both educators and students and reveals gaps in the research that helped guide the current research project. Primarily, the research showed strong evidence to look at assessments from the perspectives of the educators but scarce little research looked at students' perspectives.

Teacher Perspectives

Teacher perspectives have been studied much of the time when dealing with assessment (Cizek et al., 1995; Frey & Schmitt, 2010; McMorris & Boothroyd, 1993; Ohlsen, 2007; Stiggins & Bridgeford, 1985; Tiknaz & Sutton, 2006). For example, Tiknaz and Sutton (2006) interviewed teachers with the intent of gaining educators' views of assessments and assessment tasks in their classrooms. They concluded teachers chose what assessments worked for them based on differentiation, progression, and student enjoyment, and teachers understood the importance of assessment tasks. Young and Kim (2010) commented teachers might not view formative uses of assessment as part of their instruction. Cizek et al. (1995) found great diversity in teachers' viewpoints and practices when it came to assessments. Even Scriven (1979) pointed out teacher roles might vary on assessment. The research is plentiful when looking at teachers' perspectives on assessments in classrooms.

Other literature looked into not just teachers' views on general assessment practices but into perspectives of high-stakes testing. During the era of accountability, educators noted annual standardized tests have very limited usefulness in the classrooms, the results are not timely, and many times they do not line up with the curriculum taught (Young & Kim, 2010). Stiggins (2002) criticized once-a-year assessments because they

"are incapable of providing teachers with moment-to-moment, day-to-day information about student achievement" (p. 759). Crawford, Almond, Tindal, and Hollenbeck (2002) studied teacher perspectives on including all students in high-stakes testing and found genuine concern for students with needs and reporting to the media on the scores. Skwarchuk (2004) had 113 teachers complete a survey that gave feelings toward high-stakes tests and concluded 66.2% of all teachers surveyed agreed the tests should not be used as an assessment practice on students. Furthermore, 63.9% of teachers found the exams were a waste of government money and 51.9% believed the testing discriminated against ethnic students of color from various ethnic backgrounds (Skwarchuk, 2004). All of these studies provided insight into teacher's perspectives of high-accountability tests.

Classroom assessments have been highly studied from the viewpoint of classroom teachers. Bol, Stephenson, and O'Connell (1998) identified "relatively little research on the frequency with which teachers use various types of assessment methods and what they think about the different types of assessment" (p. 323). They went on to point out:

> Although we know a good deal about the perspectives of education theorists and researchers on assessment practices, the perceptions and attitudes of teachers toward these practices are arguably more important because change will ultimately have to occur at the classroom level. (p. 324)

They concluded teachers were prepared to develop and give their own close-ended tests but teachers had little confidence the tests were valid measures of student achievement (Bol et al., 1998). Black and Wiliam (1998b) suggested teachers could find out what students needed to know through classroom assessments by including observations, discussions, and reading student work. Tiknaz and Sutton (2006) performed semi-structured interviews and observations to better understand teachers' views. They found teachers rarely used self-assessments and peer assessments as methods of classroom

assessment and thought real assessment data were in their heads. Tiknaz and Sutton (2006) reported, “Teachers’ professional craft knowledge and beliefs had a considerable impact on teachers’ planning of assessment” (p. 340). Ohlsen (2007) found teachers preferred to make their own classroom assessments while Young and Kim (2010) determined teachers’ beliefs about conceptions of teaching influenced their assessments. Impara, Plake, and Fager (1993) studied teacher perspectives on classroom assessments and the amount of training teachers had on testing and measurement. From the 555 responses received, the researchers found most teachers were positive about classroom assessments helping their instruction, standardized tests were not useful, and rarely did teachers show zero interest in improving classroom assessment. A myriad of literature looked at the perspectives of teachers when dealing with classroom assessments. However, literature dramatically decreased when looking at classroom assessments through the eyes of students.

Student Perspectives

Both educators and researchers know students’ perspectives are important in education. Many, if not all, higher education centers identified student perspectives as important and used class surveys that relied solely on the vantage points of students (Ferguson, 2009; Gresty & Edwards-Jones, 2012; Yang, 2010). Other studies looked at teacher evaluations in relation to teaching effectiveness (Hassan, 2009; Tang, 1994). Similarly, Onwuegbuzie et al. (2007) conducted a mixed-methods study that examined students’ perspectives of effective college teachers. They concluded teacher evaluation forms are “developed theoretically” and “omit what students deem to be the most important characteristics of effective college teachers” (Onwuegbuzie et al., 2007, p.

151). Some studies investigated college-level students' viewpoints on assessments. Seymour and Chance (2010) gave surveys to students identifying student perspectives on assessment practices given by faculty. This study focused primarily on architecture and landscape architecture students and faculty who sometimes used alternative forms of assessments. Other higher education literature examined assessment practices influencing the quality of students' learning from students' perspectives (Watty, Jackson, & Yu, 2010) and also looked at pre-service teachers' perspectives of assessment practices in their education courses (Wren et al., 2009). All of these studies examined perspectives of students at the higher education level and only some looked at students' perspectives on assessments at all.

Only a small handful of studies have examined secondary students' perspectives, let alone views of classroom assessments. Most of these studies looked at what is called the "assessment environment." The assessment environment is a phrase termed by Stiggins and Conklin (1992); it is the way teachers communicate on assessments and provide feedback to students. Stiggins and Bridgeford (1985) added to the literature by studying the nature and quality of classroom assessments given by teachers. Brookhart and DeVoge (1999) deemed the term as the overall sense or meaning students make of the assessments given by teachers. It was later added that each classroom assessment is perceived by each student individually (Brookhart & Durkin, 2003). Alkharusi (2010) contributed the theory that students develop perceptions of the classroom assessment environment based on group as well as individual experiences. Of these studies on the assessment environment, Stiggins and Bridgeford (1985) surveyed teachers, Brookhart and DeVoge (1999) sampled two elementary teachers and 33 elementary students, and

Brookhart and Durkin (2003) sampled one teacher. Only Alkharusi (2010) looks at secondary students' perspectives. However, Alkharusi's (2010) study was done in Oman where public school settings are quite different from those in the United States.

Some studies focused on secondary students' perspectives but they were limited. Two such studies all came from studies done in the same country by the same researcher. Alkharusi (2010) studied the relationship between teachers' assessment practices and students' perceptions of the classroom assessment environment. In this study, 1,636 students in ninth grade participated. However, this study did not ask students to identify what assessment practices were most useful to the students and their learning. The other study, also conducted by Alkharusi (2011) in Muscat public schools in Oman, had a total of 450 tenth grade students enrolled in English language courses. The survey used in this study again looked at the assessment environment as a whole and sought out two perceived environments: learning- and performance-oriented. Again, student perspectives of what was helpful to their learning were not identified. One empirical study done by Rieg (2007) investigated perceptions of junior high school teachers and their students on the use of various classroom assessments: 119 students and 32 teachers participated from three large schools. Results from students indicated "compelling reasons to use various ways to assess… students. Many students are not adept at taking paper and pencil tests yet can demonstrate their knowledge and achievement orally or by constructing a project" (Rieg, 2007, p. 219). Many of the classroom assessments students identified as being useful to their learning were "not perceived by those same students as being used by teachers in the classroom" (Rieg, 2007, p. 220). Interestingly, participating teachers in this study concluded their choices of classroom assessments might not be

effective with their students. Conversely, the student populations of choice were students who were considered "at-risk" by the three major schools they attended. Although these studies looked at secondary students' perspectives of classroom assessment related experiences, none of the studies were done in secondary public school settings with a general population of students in the United States.

Summary

This study aimed to determine if there was a difference between secondary math students' perspectives of the assessments they were given of classroom assessments and their age, gender, or math level and to compare secondary math teachers' perspectives of their assessment practices with those students.

The review of literature articulated a thorough understanding of what assessments are in education and where gaps remain when researching secondary students' perspectives. The first section of this chapter reviewed relevant literature surrounding what assessments are and two general types of assessment practices: formative and summative. The second section of this review found literature supporting what classroom assessments are and how they are used in classrooms. The final section of this review examined research regarding both teachers' and students' perspectives and found large gaps in viewpoints from the students.

The next chapter focuses on the methodology, discusses the mixed methodology design, and gives insight as to why it was the best choice in this study. Research questions and a table showing the convergent parallel research design are also given. The desired participants, instruments, and tools used to collect data are discussed. Implementation for data collection and analysis procedures are also discussed.

CHAPTER III

METHODOLOGY

The purpose of this research was to use a convergent parallel design (Creswell & Plano-Clark, 2011) to investigate if there were differences between secondary students' perspectives of classroom assessments (Young & Kim, 2010) and their age, gender, or math level and to compare secondary math teachers' perspectives of their assessment practices with those students. This complex research problem led to the methodological approach for this study. The methodology guided the methods used for data collection, analysis, and interpretation. This chapter provides an overview of the methodology chosen and the reasons the methodology correctly fit the research problem, the justification of the design of the study, and how the data were collected and analyzed. The following research question guided this study:

Q1 Are there differences between secondary students' perspectives of classroom assessments based on their age, gender, or assessment strategy, and how do students' perspectives of classroom assessments compare with their secondary math teachers' perspectives of classroom assessment practices?

Research Methodology

Creswell (2002) defined quantitative methodology as done when the research problem involves looking at relationships among variables or looking at trends. This methodology uses statistical methods to help researchers make generalizations about a population. Quantitative data are usually gathered through surveys or questionnaires

designed to look at variables of interest. Qualitative methodology is defined as having major characteristics including understanding the meaning of the experience from participants' perspectives, the primary instrument in collecting data is the researcher, is an inductive process, and rich descriptions are needed (Merriam, 2009). Qualitative methods are appropriate when the problem involves an explanation of a phenomenon or a more detailed understanding. With qualitative methods, inferences are not made to a greater population but a more in-depth understanding of a phenomenon or participant is sought. Qualitative data are normally collected through interviews, observations, and a collection of artifacts. For the current study, the qualitative data were analyzed simultaneously with the data collection (Merriam, 2009) and coded to identify themes. The quantitative data were analyzed using statistical methods and statistical software such as SPSS and Factor. Both strands of data collection were used for this mixed methods research due to the complex nature of the question; by using both strands, the data provided a more complete understanding of the research problem.

The study was broken into four parts. The first part of the study focused on a survey for secondary math students administered during the fall 2013 semester. Specifically, the survey focused on the assessment practices secondary math students were given and if survey response differences existed between genders of the students, ages of the students, or math levels between classes of students. This part of the problem was addressed quantitatively using correlations, *T*-tests, and analysis of variance (ANOVA) tests using statistical software such as SPSS to describe the similarities and differences between variables. The second part of the research study focused on understanding the viewpoints of the student participants. Comprehending what students

were thinking when they took the survey allowed depth in understanding specific examples of the assessments and assessment strategies used in their classes. Using focus groups gave strength to the validity of the answers on the surveys and provided a more complete understanding than would otherwise have been obtained through surveys alone (Edmunds, 1999). This was addressed qualitatively. The researcher led focus groups of students who participated in the survey to explore what students thought about when completing items. Using focus groups also allowed the researcher to ask questions to students from every math classroom in the building to better conceptualize why students responded the way they did on the survey. The focus groups were digitally recorded and transcribed. Thematic coding was used to identify trends and peer review was used to see if the findings were plausible from the data collected (Merriam, 2009). The third part of the study involved understanding the viewpoints of the secondary math teachers. This was a necessity because it allowed greater depth of understanding through the use of parallel surveys and from semi-structured interviews conducted by the researcher. These provided not only the perspectives of the educators but awareness of the assessment practices teachers used in their classrooms not identified on the surveys. It also allowed clarification of the assessment strategies used by the educators. The fourth and final part of the research was focused on merging the data sets for interpretation (Creswell & Plano-Clark, 2011). It involved the researcher drawing conclusions from combining and analyzing the two strands of data. While looking to see if relationships existed, the multiple parts of the research problem being compared used different methodological approaches. Therefore, a mixed methods approach was used so all four aspects of the research question could be answered.

Study Design

In recent years, mixed methods research has grown in interest and been redefined as it continues to emerge (Creswell & Plano-Clark, 2011). In mixed methods research, the focus is to integrate both quantitative and qualitative methods and techniques to provide a more appropriate and comprehensive understanding of a research question. Johnson, Onwuegbuzie and Turner (2007) defined mixed methods:

> Mixed methods research is the type of research in which a researcher or team of researchers combines elements of quantitative and qualitative research approaches (e.g., use of qualitative and quantitative viewpoints, data collection, analysis, inference techniques) for the broad purposes of breadth and depth of understanding corroboration. (p.123)

Creswell and Plano-Clark (2011) expanded the definition more recently and described five steps necessary to design a mixed methods study including collecting both quantitative and qualitative data and integrating the two forms in some way.

Creswell and Plano-Clark (2011) suggested five steps to designing a mixed methods study. First, it should be determined that a mixed methods approach is the most appropriate to use for the purpose of the research question. In this research study, the question was complex and required not only a quantitative means to get large numbers of students' perspectives but it also required qualitative means to understand specific assessment strategies not found through a survey. Second, how the data were collected for both quantitative and qualitative methods was considered. Quantitatively, this study used a survey to gather perspectives of students on the usefulness of assessment strategies offered in the classroom. Qualitatively, focus groups of students were used to understand the thinking of answers provided on the survey using semi-structured interviews. Semi-structured interviews with teachers were also conducted to understand teacher

perspectives of their assessments, what assessment strategies were most useful to themselves as educators, and what assessment strategies they believed were most beneficial to their students. Third, the collecting of the data and the maintaining of best practices for both quantitative and qualitative methods while determining the amount of interaction between the strands were used. That is, quantitative data were collected and analyzed appropriately by administering the instruments in a standardized way and allowing proper interpretations of the results. Qualitative data were collected and analyzed appropriately as well by several means: triangulation, researcher's position, peer review, member checks, and audit trails (Merriam, 2009). The extent to which the quantitative and qualitative strands were kept independent of one another by design helped maintain transparency; the analysis of each strand was recorded. Fourth, how the quantitative and qualitative methods were mixed was considered. For this study, the data analysis from both strands was converged by the use of a matrix emphasizing the important variables from the quantitative results, the strong themes from the qualitative findings, and thus comparing them. Finally, the timing of when the two methods were mixed was determined. There are various times this can happen throughout a mixed methods study (Creswell & Plano-Clark, 2011) and this study had the mixing of the two methods being done in the analysis. This helped determine the design of the study.

As previously mentioned, the nature of the research problem gave reason for using both quantitative and qualitative approaches; thus, it was clear that a mixed methods approach was the most appropriate for the current study. Therefore, the steps required for this study are described as follows.

Choosing a Convergent Parallel Design

The research problem called for the collection of both quantitative and qualitative data. The plan was to keep the trustworthiness and integrity of the data intact and to prevent biases from entering before data could be properly mixed. Once all data were collected through both methods, the data were analyzed independently of one another and then integrated and interpreted (Creswell & Plano-Clark, 2011). This type of design is referred to as a convergent parallel design in which the two strands are independent of one another; the only point in the research process where mixing occurs is at the end to best explain the results. A procedural diagram of the convergent design helped guide the researcher in the complexity of the mixed method design (see Figure 1 for this study's procedural diagram).

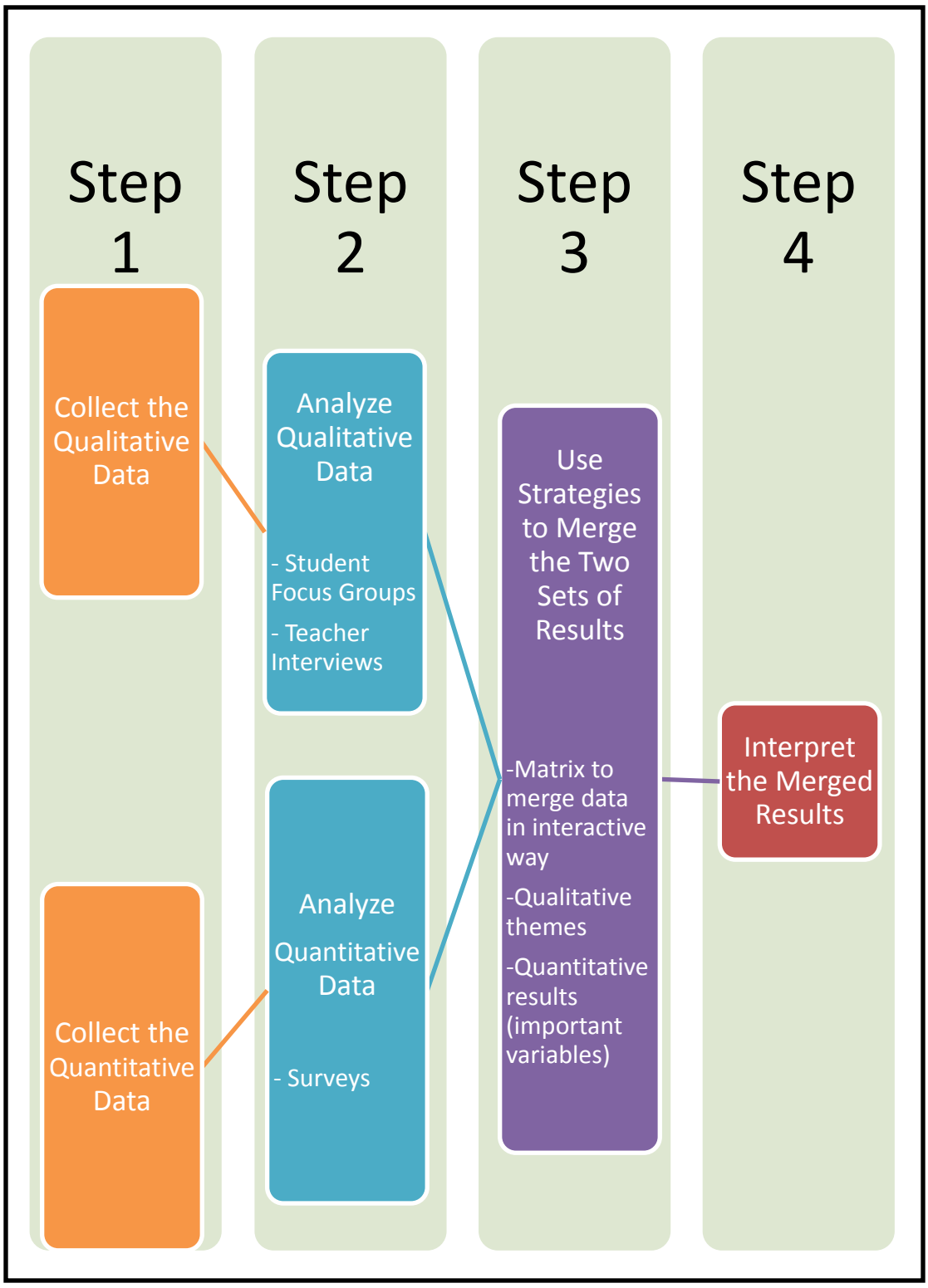

Figure 1. Procedural diagram of convergent parallel design.

Participants

Participants for this study included secondary math students (N = 420) and their high school math teachers (N = 4). All students completed the survey instrument during the fall of 2013 semester in either the months of October or November. The purpose of this analysis was to see if there were differences between secondary math students' perspectives of their classroom assessments and their age, gender, or assessment strategy. Basic demographics were collected based on gender, age, and ethnicity.

Secondary math students. The secondary math students had multiple ways of opting out of participating in the study by either form (see Appendix A) or by not filling out the survey when they were given in their classes. The final number of students who participated in the study was quite high (N = 417) with a 99% response rate. The ages of the participating secondary math students varied: the highest was 19 years of age and the youngest was 14 years-old, the average age was 16 (M = 15.64, SD = 1.44), and eight students did not share their age on the surveys. For gender, 210 males and 202 females completed the surveys, 7 individuals chose not to reveal their gender, for a total of 410 students sharing their gender (see Table 1).

Table 1

Gender of Student Participants

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Male	208	49.9	50.7	50.7
	Female	202	48.4	49.3	100.0
	Total	410	98.3	100.0	
Missing	System	7	1.7		
Total		417	100.0		

Ethnicity on the surveys was left blank for students to write in how they identified themselves. A large amount of students did not fill out this part of the demographics (*n* = 31) but of the rest of the students who did (*n* = 386) identified themselves as Caucasian (*n* = 281, 72.8%) by writing "Caucasian" or "White" or a term associated with being White for their ethnicity. The second largest group identified themselves as Hispanic or Latino (*n* = 70, 18.1%) by writing "Latino," "Hispanic," or "Mexican." The remainder identified themselves as Mixed (*n* = 20, 5.2%) by using words such as "Mixed" or "Mexican/White" to describe their ethnicity. Other (*n* = 8, 2.1%) was the term given for students who identified themselves as "Argentinian" or "Asian-Pacific" for ethnicity. Last, African American (*n* = 7, 1.8%) was the ethnicity identified by students who wrote terms such as "Black" or African-American" as seen in Table 2.

Table 2

Ethnicity of Student Participants

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Caucasian	281	67.4	72.8	72.8
	Hispanic/Latino	70	16.8	18.1	90.9
	African American	7	1.7	1.8	92.7
	Mixed	20	4.8	5.2	97.9
	Other	8	1.9	2.1	100.0
	Total	386	92.6	100.0	
Missing	System	31	7.4		
Total		417	100.0		

Teacher participants. The total number of teacher participants remained constant (n = 4) with the entire math department deciding to participate at the high school for a 100% response rate (see Appendix B for teacher consent). Teacher participants included math teachers who had been teaching for over 20 years to teachers who have been teaching fewer than 10. Each teacher participant in the study was required to develop a pseudonym for the purposes of identification and protection throughout the study. One teacher participant named himself "Diet Coke," most likely due to the large quantity of empty Diet Coke cans making a pyramid along the back wall of his classroom. Another teacher participant named herself "24" and later revealed it was a favorite jersey number. A male teacher participant simply called himself "Teacher A" while the last teacher participant simply wanted the name "X" because it "sounds cool."

There was a large amount of disparity for courses taught between the teachers in this study. For example, participant X was the only teacher who taught the college algebra class in conjunction with a nearby community college. Participant X was also the only teacher who taught AP Statistics and Statistics while other teachers such as Diet Coke taught AP Geometry and AP Calculus. A few classes were the same between teacher participants but they were primarily lower-level classes such as Geometry and Algebra II. Of the four teacher participants, the majority of the students (n = 118, 28%) had Teacher 24, while 25% (n = 106) had Teacher X, 24% (n = 103) had Diet Coke, and 21% (n = 90) had Teacher A (see Table 3).

Table 3

Teacher Participants

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Diet Coke	103	24.7	24.7	24.7
	24	118	28.3	28.3	53.0
	X	106	25.4	25.4	78.4
	Teacher A	90	21.6	21.6	100.0
	Total	417	100.0	100.0	

Survey Instrument

The instrument used in this study was created and piloted numerous times by the researcher (see Appendix C for student survey and Appendix D for teacher survey) and found to be useful. The instrument consisted of 24 items that used a 4-point frequency scale (Hakel, 1968; Simpson, 1944) about classroom assessments. Three demographic questions regarding gender, age, and ethnicity completed the instrument for a total of 27 items. The first 11 items asked questions about tests, the next 10 items asked about quizzes, and the last three asked about homework. Item 10--Tests in this class are a waste of time, Item 21--Quizzes in this class are a waste of time, and Item 23--The homework in class is a waste of time were all reverse coded due to the nature of being negative in their wording while all the other items in their sections were positive in wording. This helped to more accurately portray the meaning of the questions.

Data Collection

The mixed method approach integrated both quantitative and qualitative data into the findings via convergent parallel design (see Figure 1). In the first phase, quantitative data were collected using a 24-item survey designed by the researcher for the purposes of

this study. The second phase involved qualitative data collection by way of focus groups from the secondary math students who participated in the survey to better understand their answers on the survey. The third phase involved semi-structured interviews (Merriam, 2009) of the secondary math teachers after the first two phases were completed. The final phase involved the converging of the data during the interpretation of the strands to gather a complete understanding of assessment practices and to see where perspectives converged and diverged (Creswell & Plano-Clark, 2011).

Quantitative data collection. Quantitative data were collected using a 24-item survey from 417 participating secondary math students to see if there were differences in perspectives between age, gender, or assessment strategy. As learned from the pilot study, the instrument of choice can provide good reliability scores (see Appendix C). Data provided from this study were statistically analyzed using SPSS and other statistical software for the purpose of analysis and generalizability to a larger population. A parallel survey was provided to teacher participants as well (see Appendix D). These surveys were used for comparative purposes only since the number of teacher participants was low ($n = 4$).

Survey procedures. Participating students and the teacher in the math classes took the surveys during the regularly scheduled time period in the classroom of the participating teacher. The researcher administered the surveys to participating students in a standardized fashion to maintain consistency. The parallel survey was also given to the participating teacher at the same time as the student survey and consisted of the same number of items in the same order with changed wording to reflect the viewpoints of the participating teacher. The two demographic questions for the participating teacher

consisted of gender and number of years taught. Once surveys were completed by all participants, folders were passed around for participants and non-participants to place their surveys in for collection. Then the researcher continued onto the qualitative data collection.

Storing data. Survey data were kept in the folders in which they were collected by the researcher, placed in a secure room with the researcher's personal computer, and later secured in a locked filing cabinet in the researcher's office until the completion of the qualitative analysis by the researcher. Once data were ready to be prepared and analyzed, it was done on the researcher's password-protected personal computer. All quantitative data were recorded in aggregate form to better protect all participants. Surveys did not have any associated names of participants or locations.

Qualitative data collection. To better understand the answers student participants provided and to gain a better depth of the assessment practices used by participating teachers, two different strategies of collecting qualitative data occurred. Each form of data collection was to provide an understanding of the perspectives of the assessments being used within high school classrooms.

Focus group procedures. The first form of qualitative data collection was to have a focus group of the student participants (Merriam, 2009). This study used one focus group per mathematics class for each of the four teacher participants. Each focus group contained anywhere from three to five volunteering student participants with equal numbers of male and female students desired for each focus group. Volunteers were determined directly after the completion and collection of the survey instruments by simply asking the entire class for further participation. Student participants were asked to

further explain their answers in a secure location if they completed the survey. Student volunteers were selected by way of convenience by the researcher--the first two male students and the first two female students who volunteered were chosen. Rarely was it allowed to have more than four volunteers for the focus groups but occasionally a participating student was adamant about sharing perspectives, so exceptions were made by the researcher.

The volunteering student participants followed the researcher to a small office either in the school library or in a large storage room where the focus groups were held. All students were read the same information before the semi-structured focus groups began. Participants were encouraged but not required to participate; the researcher emphasized this at the beginning of the focus group. The digitally recorded focus groups were conducted with roughly eight questions to help lead the discussions. Questions included references to the survey student participants just took in class and asked participants to describe personal feedback in their own words and elaborate on other assessment strategies being done by the participating teacher in class that were not be covered by the survey (see Appendix E for interview questions). Upon completion of the focus group, the researcher answered questions they had and escorted participants back to their math classes.

Interview procedures. The second form of qualitative data collection was a digitally recorded, semi-structured interview (Merriam, 2009) with the participating teachers that occurred throughout the day at the convenience of the participating teachers. All interviews were digitally recorded; pseudonyms were chosen by participants and required for all participants and settings. Times included during lunches, during planning

period, and after school. Each participating teacher was interviewed with the same semi-structured format to obtain better depth of the different kinds of assessment practices they used in their math classes, which practices they felt were most important to them, and which assessment practices they felt their students valued most. These guiding questions were derived from the pilot study with critiques from interviewing participating teachers.

Storing data. Qualitative data were stored in separate folders on digital recorders by the researcher. The digital recordings were then uploaded onto the researcher's personal password-protected computer and were given a unique two-digit number. Files for each participating teacher were made using the pseudonym derived by the participating teacher during the interview. Copies of the digital recordings were then placed in appropriate files to maintain organization. Interviews were transcribed and kept with the audio recording files on the personal computer of the researcher.

Data Analysis

Data analysis occurred in three stages. The first two stages were separate analyses of the quantitative and qualitative data. The merging of both the quantitative and qualitative data followed for final analysis in this mixed method study. Qualitative data analysis occurred first to prevent the influence of finding themes through biases that would likely occur if the quantitative data were analyzed first.

Qualitative data analysis. The data analysis procedures advocated by Creswell (2007) guided the data analysis and the construction of theoretical models in this study. This process began with transcribing all data and then openly coding the transcriptions using ATLASti software (Muhr, 2014). Both focus groups and participating teacher interviews were transcribed and analyzed separately. Following open coding, thematic

coding was employed to discover premises that emerged from the student data and teacher data separately.

The original recordings and transcriptions of the interviews and focus groups were revisited before and after each round of coding to ensure the true meaning of what was said was captured accurately. No meanings were fabricated.

Quantitative data analysis. Statistical analysis of the data provided by the student surveys was done using SPSS v 16 and factor statistical software. Factor analysis, descriptive analysis, correlations, and one-way ANOVA tests (Gall, Gall, & Borg, 2007) compared groups of students.

Mixed data analysis. After the quantitative and qualitative data were analyzed separately, they were brought together for the purposes of comparing the two data sets. Creswell and Plano-Clark (2011) suggested presenting quantitative results and then qualitative findings in the form of quotes that could be easily compared. Following the data is a comment specifying how the qualitative quotes either confirmed or denied the results from the quantitative analysis.

Credibility and Trustworthiness

The researcher was affiliated with the university and, therefore, was held accountable by the same institution. An expedited Institutional Review Board application was completed and approved on July 26, 2013 with a one-year time period (see Appendix F for school consent form). All doctoral committee members had access to all data and analyses that took place in this study. Committee reviewers are respected experts in higher education and have many years of experience. Concerns brought forth regarding transcribing, coding, and recoding were addressed immediately by the researcher.

Quantitatively, the researcher used a pilot study to test the reliability scores of the instrument being used. For this study, a pilot study was used and the resulting coefficient alpha ($\alpha = .82$) yielded scores that were reliable with a respectable population ($\eta = 80$). Objectivity was another criterion present in this study; standardized conditions and fairness (Gall et al., 2007) were followed throughout the survey distribution and data collection of the surveys. To aide in this study's objectivity, the surveys were in closed-form using a Likert-type scale, were primarily self-administered, and answers on the survey were not open to subjectivity. The validity or interpretation of the data scores (Gall et al., 2007) was reviewed by experts in quantitative research in the department of statistical research at the university. The scores were also shared with the teacher participants involved in the study in April of 2014. The researcher explained the descriptive data and answered questions. This also helped increase the validity of the scores.

Qualitatively, the researcher used multiple strategies to maintain trustworthiness including techniques such as triangulation, peer reviews, member checks, audit trail, and spot checking with participants. Triangulation included the surveys given, the focus groups from student participants who took the surveys, and surveys and interviews with the participating teachers. Peer reviews were done by university supervisors and other professionals who looked at raw data to make sure the world views, biases, and relationships to the study were soundly interpreted by the researcher (Merriam, 2009). Member checks occurred in this study by having each teacher review his/her transcribed interview to ensure that the transcription and interpretation by the researcher were correct. Participants verified the accuracy of the interpretation of the data. An audit trail

was used by means of a researcher's journal that gave explicit details of the times, locations, descriptions, issues, ideas, and decisions the researcher encountered on a day-to-day basis while conducting the research in the school. Finally, spot checks were used throughout focus groups and interviews with all participants to make sure participants wanted to continue with the research.

Summary

This study aimed to determine if there was a difference between secondary math students' perspectives of the assessments they were given of classroom assessments and their age, gender, or math level and to compare secondary math teachers' perspectives of their assessment practices with those of the students. The researcher collected both quantitative and qualitative data from surveys, focus groups, and interviews. Data were then analyzed using statistical and textual methods. The integrated data analyses and findings are discussed in Chapter IV.

CHAPTER IV

RESULTS

The purpose of this study was to use a convergent parallel design (Creswell & Plano-Clark, 2011) to investigate if there were differences between secondary students' perspectives of classroom assessments (Young & Kim, 2010) based upon their age, gender, or assessment strategy and to compare secondary math teachers' perspectives of their assessment practices with those of their students. This complex research problem led to the methodological approach chosen for the study. This methodology guided the means used for data collection, analysis, and interpretation. This chapter provides an analysis of the quantitative and qualitative data collected and then a merging of the two data strands for a more complete answer to the research question. The following research question guided this study:

Q1 Are there differences between secondary students' perspectives of classroom assessments and their age, gender, or assessment strategy, and how do students' perspectives of classroom assessments compare to secondary math teachers' perspectives of their assessment practices?

The data collection was done in a school district containing 1,804 total students; 487 were in the high school where the data were collected. The high school is located in a small town with a population of roughly 4,500. According to the state website, this is the only high school in the school district. In the 2012-2013 school year, 79.34% of students were proficient or advanced on the state test scores as compared to the rest of the

state, which reported 68.72% in high schools. In the 2012-2013 school year, the high school met the state expected academic achievement and postsecondary workforce readiness scores, while it was approaching the state expected academic growth and academic growth gap scores. Enrollment continued to increase from 480 to 488 students between the 2010-2011 to 2011-2012 school years and again from 488 to 501 in the 2011-2012 and 2012-2013 school years.

The principal of the building gave approval of this dissertation when it was first introduced. He was encouraged enough to take the researcher through the school to meet the entire math department and introduced the researcher to each teacher. The principal gave quite a bit of support in the introductions and made comments such as "This will be good for all of us" and "he will be doing this here." He even went so far as to suggest to the researcher that it would be an interesting study within this particular math department.

The data were collected from the entire math department in the school and consisted of four participating math teachers using the pseudonyms of "Diet Coke," "24," "Teacher A," and "X." Moreover, the data were collected concurrently and analyzed separately. The researcher determined less bias would occur if the qualitative data were analyzed before the quantitative data to prevent preconceived themes from appearing. Thus, the data are presented in the way they were analyzed: first the qualitative data and then the quantitative data. Once both were analyzed separately, the two data strands were merged to provide a complete view.

Dan Meyer Strategy

In this study, a new strategy emerged as one of importance. It was dubbed the "Dan Meyer Strategy" because this is how teacher participant Diet Coke described the

assessment techniques he used during his interview. Upon further research, it was found that Dan Meyer currently studies math education at Stanford University, speaks internationally, and works with textbook publishing companies (Meyer, 2007). He also has a personal blog where he updates ideas on teaching math as well as thoughts on curriculum, assessment, and new ideas in math classrooms.

Qualitative Analysis

To better understand the answers student participants provided and to gain a better depth of the assessment practices used by teacher participants, two different strategies of collecting qualitative data occurred. The first section analyzes the focus groups of student participants (Edmunds, 1999). The second strategy analyzes the semi-structured interviews with the teacher participants (Merriam, 2009). Each form of data collection provided an understanding of the perspectives of the assessments being used within high school mathematics classrooms. A model was developed to better understand the themes that emerged and the relationships that existed between the themes. Figure 2 provides a visual representation of the model.

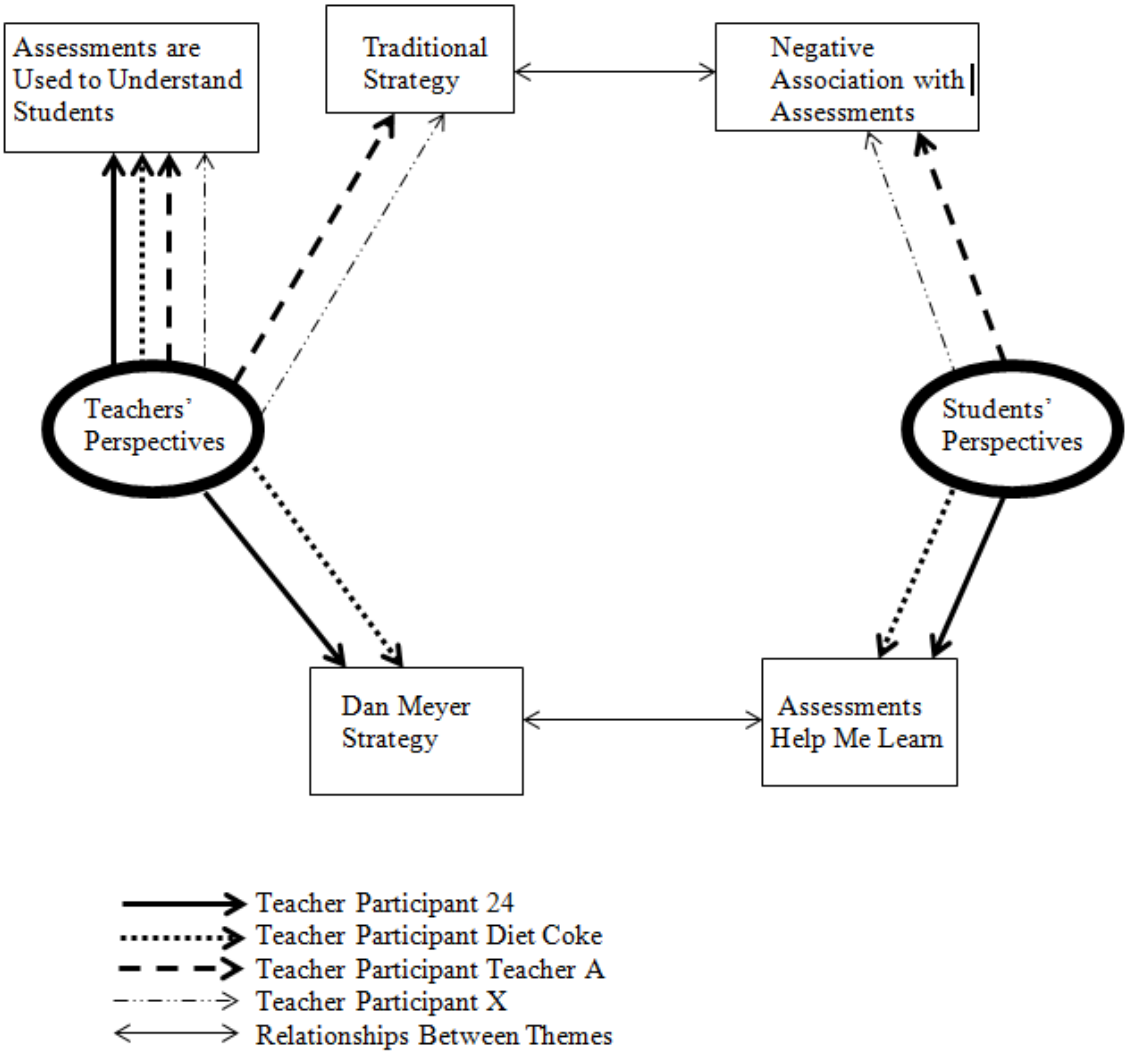

Figure 2. Visual representation of model.

Student Focus Groups

The first form of qualitative data collection was to have a semi-structured focus group of the student participants and analyze the transcripts using an open coding strategy to revisit the quotes to find themes having a reoccurring pattern (Merriam, 2009). This required all focus groups to be transcribed into Microsoft Office Word documents and then uploaded as primary documents to Atlasti (Muhr, 2014), a software program

that helped organize the qualitative data. Open coding and further revisions were used to find more specific thematic codes as they emerged. A total of 24 focus groups were conducted with an average of two male and two female students from each class. A total of 24 classes were studied. Content areas ranged from College Algebra, AP Calculus, and AP Statistics to remedial Algebra and remedial Geometry. From the focus groups, the following two major themes were identified: (a) Assessments Help Me Learn and (b) Negative Association with Assessments.

As these themes emerged from secondary math students, it is important to address important aspects of where these themes came from. The first theme of Assessments Help Me Learn only emerged from two of the four teachers' classrooms. Teacher participants who chose pseudonyms of 24 and Diet Coke consistently had this theme emerge in almost every focus group for every class. Rarely were there contradictory comments, although few existed; it became the overwhelming message of these secondary mathematics students--whatever assessment technique these participating teachers used in their classes was having a strong impact on students' perspectives across the board. Consequently, these two teacher participants also were the two teachers who used the "Dan Meyer Strategy" for giving assessments in their classes (see Figure 2).

The Dan Meyer Strategy was the name of the assessment strategy given by participant Diet Coke in his interview. In the interview with Diet Coke, the name Dan Meyer was referenced by Diet Coke as to where the assessment strategy emerged. In the interview, Diet Coke explained how it worked and the philosophies behind it. All this information was easy to access later using a general internet search. Thus, it is the name

given to the assessment strategy given weekly with rotating concepts for the duration of this dissertation (Meyer, 2007).

The opposite perspective was delivered from students who were in the classrooms of teacher participants who chose the pseudonyms Teacher A and Teacher X. These students adamantly shared views that emerged into the second theme--Negative Association with Assessments. The assessment techniques those participating teachers used in their classes were having a strong impact on their students' viewpoints and they verbally shared their opinions in a much different light than the students under the first theme. Both teacher participants X and Teacher A used the "Traditional" assessment strategy in their classrooms (see Figure 2).

The "Traditional Strategy" was the name given to the strategy used by participants Teacher A and X in this study. In the interviews with Teacher A and X, the term "pencil-and-paper" was used to describe the assessment strategies they used in class. This was further explained as "cut-and-dry, old-school test" during one of the interviews. This strategy was summarized as having periodic quizzes throughout a unit a couple of weeks apart and a large chapter test at the end of the unit covering usually a month or more of material.

Theme I: Assessments help me learn. As part of the first major theme, students expressed their attitudes and opinions toward the classroom assessment strategies used in their mathematics classes. Due to the nature of high school students, explanations were not very long but were consistent for certain teachers. For example, when focus groups were asked the question "What helps you learn in this class?" many students responded overwhelmingly with a positive view of the assessments they were given in class. One

female participant said, “For me it’s the assessments. Just getting four chances to redo it. And even if you do it right once, you have to do it right again. And he makes it harder. And it just [pause] I feel really confident until up to now I know everything.” Another male student shared, “Tests. If it wasn’t for the tests, I would be failing this class.” Other quotations included, “They’re effective because it helps me learn,” “You keep learning instead of giving up,” and “Assessments are very helpful.”

These comments were consistent for two classrooms. Teacher participants Diet Coke and 24 had overwhelming comments such as these; thus, the majority of their students created the first theme. Between the 12 focus groups for participant teachers Diet Coke and 24, this theme of assessments ‘help me learn” was mentioned over 40 times.

Theme II: Negative association with assessments. Contrary to the first theme that emerged was the second theme of negative association with assessments. This theme emerged from the other two classrooms in the study-- participants Teacher A and Teacher X. Students from the 12 focus groups from the two classrooms commented 70 times about negativity and the assessments used in class. From these comments, four subthemes emerged from the data and were later coded into the theme Negative Association with Assessments. Those four subthemes were “difficulty,” “I don’t understand where the questions come from on tests,” “failure”, and “not understanding what I did wrong.” The first subtheme of “difficulty” had comments made by students expressing the difficulty of their assessments; they included examples such as “For me the test is like everything we’ve learned all jumbled together and I have a hard time knowing if I do this for this problem or this for this problem.” The second subtheme, “I

don't understand where the questions come from on tests," had quotations from students supporting the idea of not understanding where the test questions came from: "We did homework on certain things and we have like 10 other things on the test and I don't understand where they came from." Another supporting example of not understanding the test questions included this particular comment made with great frustration by a student:

> That's why I don't understand where the tests come from because like [pause] if you wanted to you could do the homework like in five minutes before it was due and understand everything that was going on. Then you get the test and it is like, "Where did this problem come from?" It [homework] doesn't really correlate to the test as much as it should.

The third subtheme of "failure" emerged when questions asked students what came to mind in math class when the word "test" was used. Multiple quotations supported this subtheme: "Hell," "I'm going to fail this. Really, no, I'm going to fail this," "I don't ever feel prepared on every test," "Automatically going to fail," "Failing," "I have not passed one of his tests," "I passed one, but it was luck. I guessed on like every single question," "We're all failing. It's not helping at all," "My tests look like scribbles," "Just red marks all over," "He's surprising us," and finally "Your grade is going to drop."

The fourth and final subtheme of "not understanding what I did wrong" was supported by comments by students including examples such as "He doesn't really show what you did wrong. Just correct…" and "He doesn't show us what we did wrong." These comments were directly related to the feedback not only on the assessments themselves but also the feedback covered in class after assessments were returned to students. Students looked for feedback in multiple areas in class and seemed to feel they never received it.

The four subthemes of “difficulty,” “I don’t understand where the questions come from on tests,” “failure”, and “not understanding what I did wrong” combined to create the second theme of negative associations with assessments. All the subthemes had a slightly different perspective of the assessments but certainly had negative connotations linking them together.

The two themes from the math students were related to the assessment techniques the participating educators used in their classrooms. The theme of assessments “help me learn” clearly became related to the Dan Meyer Strategy and the participating teachers who used this assessment technique. The opposite was also true; the theme of negative association with assessments was visibly related to the “Traditional Strategy” and the participating teachers who used this assessment technique (see Figure 2).

Teacher Interviews

The second form of qualitative data collection included four semi-structured interviews (Merriam, 2009) with each of the participating teachers. These interviews took place throughout the day while data were being collected, primarily during planning periods or over the lunch hour ranging from 7 to 19 minutes in length. These interviews found a few things in common but because the number of participants was small, it was hard to say if these were themes. However, the two major similarities discussed fell loosely into two categories: (a) assessments are primarily used to understand students and (b) differences in assessment techniques. All participating teachers identified with the first similarity. During interviews, each teacher identified the purpose of his/her assessments was to help understand what students knew. They all agreed with the philosophy that their assessments provided opportunities to learn about their students.

Similarity I: Assessments are used to understand. The participating mathematics teachers in this study all had similar perspectives about the purposes of assessments in their classrooms. All four educators viewed assessments as a primary means of collecting data to determine what their students understood in class. For example is this quote from Participant X when referring to giving quizzes in class:

> To me it's more of a checkpoint so that I can make sure that everybody has the skills from the section of the chapter that is more recent. Make sure they have those skills up to par before we push through the rest of the chapter and if they do great, we push through. If not, it tells me okay, we need to go back and review a little bit, maybe do some test corrections or a retake, or something to get those skills up so I don't just hand them a quiz, they take the quiz, and we move on whether they're ready or not. It's a nice checkpoint.

Participant 24 also shared this view:

> To see if the kids understand. I mean that's the whole point why we're here. To teach them and they learn something. And so if they're not learning, then obviously we're doing something wrong. So we want to be sure that they're taking a little bit away from each day or week that we're doing these.

The participant identified as Teacher A identified numerous paper-and-pencil assessment strategies that were helpful in the classroom:

> Paper-and-pencil quizzes and tests because I see the students' work and I can assess how well they are processing from one step to another as they solve a problem. I can get a little bit more inside into their thought process, where it is they have gaps from previous learning. So paper-and-pencil assessments I think are the most insightful. It's not the most time sensitive because obviously there is a little bit of a lag there, but paper-and-pencil.

Finally Participant Diet Coke briefly summarized his perspectives of how assessments were useful to him when he said, "I personally can read those numbers off and know exactly what a student does and doesn't know how to do."

Overall, all four participating teachers in this study shared reflections that revealed similar views on how their assessments were used in their own classrooms to

better understand what their students were learning and understanding. These findings supported the ideas of Frey and Schmitt (2010) and Ohlsen (2007) that classroom assessments are used to inform and guide teachers.

Similarity II: Difference in assessment techniques. Although all the educators agreed on how assessments in their classrooms were used, the techniques in which assessments were given varied greatly and were starkly different. Two strategies were identified as the data were collected; the Dan Meyer Strategy and the Traditional Strategy were the two assessment techniques in which classroom assessments were given to students. The disparities of these assessment strategies led to the second similarity emerging and were unmistakably related back to the perspectives of the students (see Figure 2).

The first assessment technique was the Dan Meyer Strategy. With the Dan Meyer Strategy, the teacher participants had a unique way of assessing students. Teachers assessed students over four concepts every week; with every assessment, a new concept was added based on what students were learning and an old concept was removed. There was no official title of the assessment: the labels “quiz” and “test” did not exist in these classrooms. Only the word “assessment” was familiar to these students. With the Dan Meyer Strategy, the philosophy was to give repeated opportunities to retest a given concept with the concept getting more and more difficult every week the students saw them. It also required students to do the concept perfectly on any two of the four weeks they saw it in order to receive the title of “mastery” of that concept and thus receive full credit from the teacher (Meyer, 2007).

Two teacher participants, Diet Coke and 24, had comparable assessment strategies they believed worked. The strategy they followed is referred to as the Dan Meyer Strategy. This strategy is best described by Diet Coke with some personal justification:

> Right now I give a weekly assessment that is four problems. Those four problems rotate every week. And then the numbering on them is very specific. The first week in class I gave them four assessment items and those were intended to be review. And they were numbered one, two, three, four, then two, three, four, five, then three, four, five, six and so on. And so what that does it gives every student with the exception of the first couple of assessment items, it gives every student four chances to see an item. So if they don't get it right the first time, we go over it in class and they can figure out what they're doing wrong. And then for the next assessment take a similar problem and hopefully get it right. The caveat there is that they get so many opportunities that I make them get it right at least twice before I tell them they're done with that problem. So if on two occasions they get a right answer, they get a five out of five on the grade book. The items are broken down question by question in the grade book. I did originally do the big test, right? The end of the chapter: a month of teaching. Now let's see how much you can spit back onto the page for an hour and a half. I did those for a while. I felt like they had pretty mixed results, which is why I switched to this. Before I'd see a test in the gradebook that said 80 out of 100, and 80 out of 100 didn't really mean much to me. Now when I look at the gradebook, I see problem number one… integer operations and that kid's got a three out of five, and so I know they're making some minor mistake that I see a lot that needs to be corrected because the three out of the five means that to me. The four out of five means they've shown it to me once out of the times that they've done it. They've shown it to me correct once. And they need to show me the correct answer one more time to get the fifth point in the gradebook. So it's kind of a four plus four equals five kind-of-thing in my math class.

An example of a concept list was collected as an artifact (see Appendix G). Participant 24 used the same Dan Meyer assessment strategy in class and also justified it in a similar way:

> The main thing is I test weekly. They get four concepts a week. And we add a new concept every week and we take a concept every week. So three repeated concepts every week, one new. Basically the reason I do that is for repetition. It allows the kid to test over the same concept over a month instead of getting all this information for a chapter test, trying to remember that after a month or a month and a half and then test over it, so that's why it makes the most sense.

The second assessment technique was the “Traditional Strategy” and it was quite contrary to the first assessment practice. With this second technique, the assessments were spread out with a large accumulating summative test at the end of a unit with smaller checkpoints or quizzes along the way. Participant X and Teacher A conducted assessments in their own classes using the Traditional Strategy. Teacher A said, “Maybe once every week and a half, and then once every three or four weeks for tests” as the technique suggested. Participant X also made claims of similar Traditional assessment strategies: “In terms of formal assessments, I usually give one quiz about half way through a chapter and then one test at the end of the chapter: kind of cut-and-dry old-school tests. Here’s a test, here’s a pencil, go.”

This second similarity emerged rather clearly as differences between the assessment techniques used by the mathematics department at this particular high school. The second similarity of differing assessment techniques also emerged as being related to the perspectives of the secondary math students. This implied the differing assessment techniques were in relationship to the perspectives of the secondary math students who took them (see Figure 2).

Relationships Between Perspectives

Relationships between the perspectives of secondary math students and their teachers appeared as the data were analyzed. The two assessment strategies used in the building gave polar results from the perspectives of the students. The teachers who used the Traditional Strategy were participants Teacher A and X; regardless of the math ability of their students, the theme Negative Association with Assessments emerged as being related to the Traditional Strategy assessments. In short, if the teachers used the

Traditional Strategy of giving assessments in their classroom, students viewed those assessments negatively (see Figure 2). The teachers who used the Dan Meyer Strategy were teacher participants 24 and Diet Coke. Regardless of the math ability of their students, the theme Assessments Help Me Learn emerged as being related to the Dan Meyer Strategy of giving assessments. If teachers used the Dan Meyer Strategy of giving assessments in their classrooms, students viewed those assessments positively (see Figure 2).

Quantitative Analysis

Factor Analysis

Very little data were missing in this study, which helped carry out the factor analysis. Three separate factor analyses were conducted: one for the test items, one for the quiz items, and one for the homework items. The factor analyses were conducted to see if all the items within each set were measuring the same thing. The number of dimensions, factors, tells us how many scales can be constructed from a set of items.

The method for extracting the factor was the unweighted least squares and the number of factors was determined by parallel analyses (Timmerman & Lorenzo-Seva, 2011). Parallel analysis has been found to be the most accurate method for determining the number of factors. In sum, the factor analyses found one factor for the test items and one for the quiz items (see Tables 4 and 5).

Table 4

Factor Analysis of Test Items

Items	Factor Loadings
Item 1--The tests in class help me learn	.792
Item 2--The tests in class are useful to me	.772
Item 3--The tests in class help me understand what I have learned	.697
Item 4--The tests in class help me understand where I make mistakes	.603
Item 5--My teacher uses tests in class to give me general feedback on my learning	.517
Item 6--When I get a test back I know where I made my mistake	.533
Item 7--When I get a test back I know how to do better next time	.647
Item 8--I get an opportunity to relearn information on a test when it is returned to me	.610
Item 9--I get an opportunity to retake tests	.393
Item 10--Tests in this class are a waste of time	-.601
Item 11--I like this teacher	.420

Table 5

Factor Analysis of Quiz Items

Items	Factor Loadings
Item 12--The quizzes in class help me learn	.805
Item 13--The quizzes in class are useful to me	.833
Item 14--The quizzes in class help me understand what I have learned	.830
Item 15--The quizzes in class help me understand where I make mistakes	.778
Item 16--My teacher uses quizzes in class to give me general feedback on my learning	.662
Item 17--When I get a quiz back I know where I made my mistake	.659
Item 18--When I get a quiz back I know how to do better next time	.753
Item 19--I get an opportunity to relearn information on a quiz when it is returned to me	.694
Item 20--I get an opportunity to retake quizzes	.441
Item 21--Quizzes in this class are a waste of time	-.459

Scale Construction and Reliability

Scales were created through the summation of all the test items (see Table 4) in order to clearly run descriptive tests and analysis of variance tests. The same was true for the quiz factor items: all items needed to be added together for analysis to take place (see Table 5). Last, a total for the three homework items was also created for comparison and descriptive purposes. The correlations among the scales are presented in Table 6.

Table 6

Correlations Among Homework Items

	Item 22	Item 23	Item 24
Item 22	1.000	-.414**	.563**
Item 23	-.414**	1.000	-.291**
Item 24	.563**	-.291**	1.000

Note. Item 22--The homework in class helps me learn. Item 23--The homework in class is a waste of time, Item 24--The homework in class is used to give me feedback on my learning
** Correlation is significant at the 0.01 level (2-tailed).

Cronbach's alpha was used to examine the internal consistency of the scales. The test items had a high reliability ($\alpha = .85$) with minimal improvement if any items were deleted. The quiz items also had high reliability ($\alpha = .90$) with minimal improvements if any items were deleted. There were only three homework items and those had a lower reliability score ($\alpha = .69$) when they were grouped separately.

Correlations Among the Scales

The correlations among the scales were all statistically significant. The correlations among the test, quiz and homework scales are presented in Table 7. The correlations between the quiz and test scales were high and their correlations with the homework scale were lower.

Table 7

Correlations Among the Test, Quiz, and Homework Scales

	Test	Quiz	Homework
Test	1.000	.846**	.242**
Quiz	.846 **	1.000	.200**
Homework	.242**	.200**	1.000

** Correlation is significant at the 0.01 level (2-tailed).

Of particular note, Homework Item 22--The homework in class helps me learn had a strong positive correlation with Homework Item 21--Quizzes in this class are a waste of time (r = .56). Items 21 and 22 had very high positive correlation with Item 10--Tests in this class are a waste of time (r = .59). Students who perceived the homework in class as helping them learn also perceived the homework was used to give feedback on their learning. Students who thought homework in class was a waste of time also believed the tests in their math class were a waste of their time.

Correlation Among Demographic Variables

Very opposite relationships existed when looking at both the demographic and descriptive items. Correlations on demographic items on the survey showed very weak associations. For example, very weak relationships with gender (r < .16), age (r < .24) and ethnicity (r < – .11) existed across all items. On the contrary, extremely strong positive relationships existed for descriptive items such as the teacher students had and the assessment strategy used by the teacher (r = .86).

Relationships Between the Test and Quiz Scales with the Demographic and Instructional Variables

Age. The reported range was seven years difference (min = 12, max = 19) with only one participant writing in the age of 12. The average age of participants was 16-years-old (M = 15.64, SD = 1.44). A one-way analysis of variance (ANOVA) test was conducted to evaluate the relationship between the ages of the secondary math students and their perspectives of how they viewed their assessment strategies in their classroom. The independent variable was the ages of the students while the dependent variable was the total of the scores from the test factor items (Items 1-11) on the survey. There was no significance between the secondary math students' perspectives of their tests and their age groups, $F(7, 395) = 1.76$, $p = .10$. Thus, maturity did not have any influence on the perspectives of the students for the test items. The same was true of the quiz factor items (Items 12-21) and the students' perspectives, $F(7, 392) = 1.03$, $p = .41$.

Gender. One question guiding this study was "Is there a difference between secondary math students' perspectives of their classroom assessments and their gender?" To better answer this question, a one-way ANOVA was conducted between test factor items on the survey and the gender of the students. A total of 206 male students' scores (M = 32.16, SD = 6.69) and 198 female students' scores (M = 33.68, SD = 6.82) were compared for the test factor items. The analysis revealed the perspectives of classroom assessments differed between male and female students, $F(1, 402) = 5.15$, $p = .02$. The same differences in perspectives existed when analyzing the quiz factor items and the perspectives of males (n = 202, M = 28.46, SD = 6.78) and females (n = 199, M = 30.38, SD = 6.79), finding a significance, $F(1, 399) = 8.04$, $p = .01$. Further analysis of the

survey items identified which questions showed gender differences. Test factor Items 1, 4, 7, 8, and 10 were all significant and quiz factor items 12, 14, 15, 18, 19, and 21 all showed differences between genders (see Tables 8 and 9).

Table 8

Analysis of Variance for Test Items

		Sum of Squares	*df*	Mean Square	*F*	Sig.
The tests in class help me learn	Between Groups	4.418	1	4.418	4.891	.028*
	Within Groups	368.560	408	.903		
The tests in class are useful to me	Between Groups	1.090	1	1.090	1.151	.284
	Within Groups	386.571	408	.947		
The tests in class help me understand what I have learned	Between Groups	1.077	1	1.077	1.299	.255
	Within Groups	335.861	405	.829		
The tests in class help me understand where I make mistakes	Between Groups	4.958	1	4.958	5.991	.015*
	Within Groups	337.686	408	.828		
My teacher uses tests in class to give me general feedback on my learning	Between Groups	.073	1	.073	.092	.761
	Within Groups	320.584	408	.786		
When I get a test back I know where I made my mistake	Between Groups	.355	1	.355	.407	.524
	Within Groups	355.606	408	.872		
When I get a test back I know how to do better next time	Between Groups	5.268	1	5.268	6.024	.015*
	Within Groups	355.063	406	.875		
I get an opportunity to relearn information on a test when it is returned to me	Between Groups	5.552	1	5.552	6.066	.014*
	Within Groups	373.472	408	.915		
I get an opportunity to retake tests	Between Groups	1.105	1	1.105	.804	.370
	Within Groups	559.345	407	1.374		
I like this teacher	Between Groups	.349	1	.349	.308	.579
	Within Groups	461.255	407	1.133		
Item10New	Between Groups	7.895	1	7.895	9.805	.002*
	Within Groups	328.517	408	.805		

* Indicates significance at $p < .05$.

Table 9

Analysis of Variance for Quiz Items

		Sum of Squares	*df*	Mean Square	*F*	Sig.
The quizzes in class help me learn	Between Groups	4.026	1	4.026	4.769	.030*
	Within Groups	344.423	408	.844		
The quizzes in class are useful to me	Between Groups	2.305	1	2.305	2.494	.115
	Within Groups	376.223	407	.924		
The quizzes in class help me understand what I have learned	Between Groups	4.422	1	4.422	5.280	.022*
	Within Groups	340.859	407	.837		
The quizzes in class help me understand where I make mistakes	Between Groups	4.309	1	4.309	5.466	.020*
	Within Groups	321.681	408	.788		
My teacher uses quizzes in class to give me general feedback on my learning	Between Groups	1.938	1	1.938	2.574	.109
	Within Groups	305.709	406	.753		
When I get a quiz back I know where I made my mistake	Between Groups	1.698	1	1.698	2.144	.144
	Within Groups	322.434	407	.792		
When I get a quiz back I know how to do better next time	Between Groups	7.367	1	7.367	8.650	.003*
	Within Groups	344.957	405	.852		
I get an opportunity to relearn information on a quiz when it is returned to me	Between Groups	6.860	1	6.860	7.316	.007*
	Within Groups	378.837	404	.938		
I get an opportunity to retake quizzes	Between Groups	4.446	1	4.446	3.110	.079
	Within Groups	577.537	404	1.430		
Item21New	Between Groups	3.648	1	3.648	4.192	.041*
	Within Groups	353.291	406	.870		

* Indicates significance at $p < .05$

Assessment strategy. One major question that drove this study was to see if there were differences between secondary math students' perspectives of their assessment practices and the assessment strategies their teachers used in class. Secondary math

teachers in this study used one of two strategies: the first was identified as the Dan Meyer Strategy and the second was the Traditional Strategy. The Dan Meyer Strategy had a small majority of students (n = 217, M = 36.38, SD = 5.35) compared to the Traditional strategy (n = 192, M = 28.27, SD = 6.01). A one-way ANOVA was used to see if differences in these assessment strategies existed for test factor items. Data analysis revealed major differences in student perspectives between the assessment strategies, $F(1, 407) = 178.62$, $p = .00$. Individual one-way ANOVA tests were run on each test factor item and revealed major differences persisted at a $p = .00$ significance on all 11 test items. To further compare assessment strategies, another one-way ANOVA test was run to see if differences persisted between students' perspectives of the assessment strategies and the quiz factor items. Again, the Dan Meyer Strategy had more students (n = 217, M = 32.48, SD = 5.94) than the Traditional Strategy (n = 189, M = 25.75, SD = 6.09). Major differences persisted, $F(1, 404) = 126.74$, $p = .00$, revealing diverse student opinions existed between the assessment strategies. Individual one-way ANOVA tests were run on all quiz factor items to determine if significance levels remained consistently large. All 10 items had major significance at the same level, $p = .00$.

Ethnicity. Secondary students self-reported their ethnicity for this study. It was determined it would be important to see if differences existed between these groups of students' perspectives and both the test factor and quiz factor items on the survey. Proper data analysis required ethnicity be coded in two ways--Caucasian and non-Caucasian--in order for the means of the two groups to be compared for both the test factor and quiz factor items (see Tables 10 and 11).

Table 10

Analysis of Variance on Test Factor

	Sum of Squares	*df*	Mean Square	*F*	Sig.
Between Groups	46.181	1	46.181	1.009	.316
Within Groups	17347.777	379	45.772		

Table 11

Analysis of Variance on Quiz Factor

	Sum of Squares	*df*	Mean Square	*F*	Sig.
Between Groups	18.028	1	18.028	.385	.535
Within Groups	17604.377	376	46.820		

Two one-way ANOVA tests were conducted--one for the test factor items and the other for the quiz factor items. The test factor ANOVA did not show any significance between students' self-identified ethnicity and their perspectives, $F(1, 379) = 1.01$, $p = .32$. The quiz factor ANOVA revealed similar results, $F(1, 376) = 0.39$, $p = .54$. There did not appear to be any major differences between perspectives of classroom assessments and the ethnicity of secondary math students.

Teacher. Secondary mathematics teachers use different assessment, grading, and feedback techniques in their classrooms. In this study, teacher participants claimed the ability level of the students did not change these techniques. Therefore, it was important to explore if differences existed between secondary math students' perspectives and the

teacher who taught them mathematics. The descriptive statistics for each teacher can be found in Table 12 below:

Table 12

Descriptive Statistics for Test Factor

	N	Mean	SD	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Diet Coke	102	35.98	5.709	.565	34.86	37.10	20	44
24	115	36.73	5.013	.467	35.80	37.66	22	44
X	105	28.49	5.924	.578	27.34	29.63	14	41
Teacher A	87	29.33	6.113	.655	28.03	30.64	17	41
Total	409	32.85	6.793	.336	32.19	33.51	14	44

A one-way ANOVA was conducted to see if differences existed between the secondary math students and who taught their math course on the test factor items. Major differences did exist, $F(3, 405) = 60.21$, $p = .00$, indicating students had different perspectives on their classroom assessments based on who their teacher was. A Scheffe post-hoc test revealed similarities between teacher participants Diet Coke and 24 with a significant *p* value ($p = .81$) while also showing differences from Teacher A ($p = .00$) and Teacher X ($p = .00$). The test also showed no significance between Teacher A and Teacher X ($p = .79$). Teacher participants Diet Coke and 24 were grouped together and Teacher A and Teacher X were grouped together on all test factor items showing students had very different perspectives on the tests based on their teachers.

A second one-way ANOVA was conducted to see if differences between students' perspectives and their math teacher existed on the quiz factor items. Once again, major differences existed between the students' perspectives and who their math

teacher was, $F(3, 402) = 45.30$, $p = .00$ on all quiz factor items on the survey. A Scheffe post-hoc test was run to find where differences occurred between teachers on the quiz factor items. The test revealed a familiar output that supported strong similarities between Diet Coke and 24 ($p = .92$) and differences from Teacher A ($p = .00$) and Teacher X ($p = .00$). The test found similarities between Teacher A and Teacher X as well ($p = .08$) although the similarities were not as strong. Overall, students' perspectives of their quizzes were based on the teachers who taught them.

CHAPTER V

DISCUSSION

In this convergent mixed methods study, data were converged and analyzed to investigate if there were differences between secondary students' perspectives of classroom assessments and their age, gender, or assessment strategy and to compare secondary math teachers' perspectives of their assessment practices with those of their students. This chapter discusses the findings of this dissertation research along with implications. Then limitations and future research are shared.

Purpose

The purpose of this study was to see if differences exist between secondary math students' perspectives of their classroom assessments and their age, gender, or assessment strategy and to compare students' perspectives with those of their secondary math students. A convergent parallel design was used to gather concurrent qualitative and quantitative data. Creswell and Plano-Clark (2011) claimed, "They are necessary in a mixed methods study because both quantitative and qualitative data collection are central to this form of inquiry" (p. 162). The quantitative portion of the mixed methods question focused on differences between students' age, gender, and assessment strategy and their perspectives of the classroom assessments they were given. The quantitative data collection occurred by way of administering surveys to secondary math students. The qualitative aspect of the mixed methods question was to see how students'

perspectives of their classroom assessments compared to the views of the math teachers who gave them the assessments. Qualitative data were collected by focus groups of students and interviews with secondary math teachers. Both quantitative and qualitative strands were collected concurrently and analyzed separately. With the combining of the two data strands, it was possible to get a more complete analysis due to the complexity of the research question (Creswell & Plano-Clark, 2011).

Before converging the data strands, a crucial clarification must be made. The initial research question at the beginning of this study had the intent of determining which assessment strategies secondary math students favored when comparing tests and quizzes along with other unique things teachers did to assess students. There was no prior knowledge of the actual classroom assessment strategy labeled the Dan Meyer Strategy (Meyer, 2007) before this research took place by the researcher. While collecting data, the Dan Meyer Strategy emerged as its own unique philosophy where both students and teachers had distinctive perspectives about it. So the research question took on a slightly different meaning. Thus, this dissertation research emerged to allow comparisons between two assessment strategies in high school mathematics classrooms: the Dan Meyer Strategy (Meyer, 2007) and the Traditional Strategy.

This chapter examines the converged data strands and the perspectives each offers to better understand the research question that guided this study. Since age was not found to be a significant aspect of students' perspectives, it will not be in the discussion. Each of the following sections combines qualitative and quantitative strands to evaluate the data through the differences between two classroom assessment strategies, differences in

students' perspectives between teachers, and differences between the genders of the students.

Differences in Assessment Strategy

There were major differences in the two assessment strategies in this doctoral study according to the teachers and the math students. Due to the nature of the Dan Meyer Strategy, tests and quizzes were one and the same. They were only known as "assessments" in the classrooms where this strategy was being used. When referring to the two sections on the survey given by this researcher, one student in a focus group explained how both sections were essentially the same from her perspective: "On the tests [survey] how it shows quizzes and tests, we just basically have one assessment per week. It's not different." On the contrary, the Traditional Strategy offered periodic quizzes followed by large chapter tests. Teacher participant X explained: "I usually give one quiz about half way through a chapter and then one test at the end of the chapter."

The questions used on the Dan Meyer Strategy versus the Traditional Strategy were very different. Students who were assessed using the Dan Meyer Strategy only had four concepts they had to perform every week. The new concept was created by the teacher based on the standards and what was taught in class during the week. This helped guide what concepts were being assessed. So the questions were not necessarily preconceived weeks in advance. Many times, they were created the day before the assessment or if they had been designed ahead of time, they were tweaked to make sure the assessment covered a concept in a narrow fashion.

The questions on the Traditional Strategy had numerous concepts to perform based on the material taught in class. For example, some quizzes might only cover three

concepts and the next quiz might cover five concepts. The chapter test could cover as many as 9 or 20 concepts depending on the unit. The Traditional Strategy questions were distinct and based on what the teachers selected as important form the chapter, similar to the Dan Meyer Strategy. However, many times, these questions were taken directly from books or were required to be given via the curriculum.

So at their very core, these strategies were very unique and not very similar. As seen in Table 13, students had polarizing views of these assessment strategies along with their teachers.

Table 13

Differences Between Assessment Strategies

	Dan Meyer Strategy *M* = 2.94, *SD* = .88	Traditional Strategy *M* = 2.19, *SD* = .88
A teacher's perspective	"I use them as a learning tool."	"A necessary evil that I need to make sure."
A student's perspective	"I feel really confident up until now I know everything."	"I'm going to fail this. No, really, I'm going to fail this."

Assessment Strategy & Test Factor Items – $F(1, 407) = 178.62$, $p = .00$
Assessment Strategy & Quiz Factor Items – $F(1, 404) = 126.74$, $p = .00$

Views of the Students

Students as a whole felt very different about these two assessment strategies in their classroom on both the test factor items, $F(1,407) = 178.62$, $p = .00$, and the quiz factor items, $F(1, 404) = 126.74$, $p = .00$, showing significant differences in how students viewed the tests and quizzes in their classrooms. These differences seemed to imply the students in this study favored the Dan Meyer Strategy over the Traditional Strategy. As

one student stated, “It’s a very easy format. This is only my second year learning under [name omitted] and I much prefer his testing method than to [name omitted].” The quotations by students (see Table 13) generally summed up the opposing views--“I feel really confident up until now I know everything” and was a strong over-arching theme for how students generally felt using this method. “I’m going to fail this” was a more consistent overarching theme for students under the Traditional Strategy. The reasons for the disparities could be due to the philosophical differences of timing between them, the difficulty increase for each concept, the number of concepts each strategy covered when given, and differences in how the assessments were graded.

First, the Dan Meyer Strategy offered a unique style of timing that allowed students to see the same concept four different times with each being slightly more difficult than the previous. So they saw the same concept repeatedly, therein seeing the concept more often. This also implied the students saw a concept for four weeks in a row. In contrast, the Traditional Strategy was limited in its timing. Using this strategy, one could go over a concept on a quiz and then have multiple weeks before being tested on that concept again. During the time between the test and quiz, students would likely be learning new information and not have multiple opportunities to be exposed to the concept they would be tested on, thus seeing the concept less. Students supported this implication with comments about the Dan Meyer strategy: “Short and simple. Like it’s not very long. It’s just a 5 minute test” and “We don’t have to do this big study guide and it turns into a week-long project just to take one test.” One student under the Traditional Strategy made a comment about the amount of time between concepts: “We’re not going to remember them anyways to be honest.”

Further explanation could be the difficulty jump in concepts from the time they are introduced to the time students are tested over the concepts. The Dan Meyer Strategy builds, i.e., students are introduced to a concept and as they continue to see the concept every week, it gets slightly more difficult. At its core, the Dan Meyer Strategy is repeated testing and, through that, repeated teaching. Participant Diet Coke claimed:

> Test anxiety is something that doesn't exist with my assessments… It eliminates that stress. So a lot of my students do, I think, look at is as a way: a learning tool as much as a way for me to figure out what they do and don't know. And I wouldn't change back to the big test for anything in the world after seeing that kind of a change in my students.

This allows students to have little to no test anxiety because students know they will get another opportunity to be successful the next week plus when their teacher goes over the assessment concepts in class, students get to relearn. One comment by a student supported this: "It's really helpful because we basically go over the ideas we had in class and only maybe a couple of problems for each idea and it just really helps to understand what is going on." The Traditional Strategy does not give students as many opportunities to see the same concept; therein, the difficulty level of the concepts is likely to increase dramatically more for each exposure. Most likely, students will learn a concept in a day or two, get quizzed on it in a slightly more difficult way a few weeks later, and then have to know it to show mastery on the test, which could be many weeks later. At best, it would seem students do not have as many opportunities to learn the concept, let alone be successful on the concept when it counts. Students have to recall the information on their own with less exposure and understand it at the most complex level with fewer opportunities to learn. As one student put it, "It's like going from this level to that level."

Another possible explanation for the differences between the students' perspectives of the assessment strategies could be because of the differences in the number of concepts on the assessments themselves. The Dan Meyer Strategy only allowed four concepts to be present on any given week, whereas it seemed as though the Traditional Strategy covered many more concepts since tests were over entire chapters spanning multiple concepts. Students who were assessed under the Dan Meyer Strategy shared insights into the benefits of only having four concepts on each assessment: "It's not a lot of pressure" and "You keep learning instead of giving up." When asked how many concepts were covered on quizzes, one student under the Traditional strategy identified "generally two or three" but continued on about the difficulties of recall between the quizzes and tests:

> Because it's like 3.1 through 3.9 or something, the quiz is going to be after like 3.3, 3.6, 3.9, so we take three separate quizzes all on the stuff that we just learned. And the test are you have to remember everything from 3.1 to 3.9, and being able to pull it and use it, especially if you haven't learned it in a month.

The last potential explanation of why the differences between the assessment strategies existed could be because of how the assessments were graded. Grading is very controversial in education (Carey & Carifio, 2012); how tests are graded by teachers likely has an influence on how students feel toward those tests. The Dan Meyer Strategy has a unique system of grading using a 5-point rubric and requires perfect answers on a concept multiple times before the grade gets bumped up to mastery or a 5. It does not allow students to get lucky just once and get a high score. It forces them to be successful two times but the two times can come in any order. Teacher participant Diet Coke said, "The caveat there is that they get so many opportunities that I make them get it right at least twice before I tell them they're done with that problem. So if on two occasions they

get a right answer, they get a five out of five on the grade book." One student under the Dan Meyer Strategy felt the style by which he was graded helped students: "It also helps kids; it helps them feel better because they get a good grade." So there was evidence to suggest that the Dan Meyer Strategy's unique grading system impacted students' perspectives.

Views of the Teachers

Some evidence suggested the teacher participants viewed their own assessment strategies differently due to their own philosophies of assessment. However, caution should be used when looking at the perspectives of the educators since so few teachers were involved in the study. The quotations from Table 13 showed two possible philosophical views based on the assessment strategy used in the classroom. "I use them as a learning tool" was a very unique perspective that carried strongly with the two teachers who used the Dan Meyer Strategy. It was the constant interaction with the concepts that potentially helped mold this perspective. Since students had multiple encounters with a concept, the teachers went over the concepts with their classes every week. This gave multiple opportunities for teaching and re-teaching while also giving numerous chances for students to be successful. Hence, the concepts philosophically directed teachers using them to better see how assessments could be more useful in the classroom to both teachers and students; hence, they were learning tools.

A very opposite view was shared by one of the teachers who used the Traditional Strategy. "A necessary evil that I need to make sure… here's a test and show me where you're at" was a very general theme used to summarize the perspectives of the teachers who used the Traditional Strategy. The mentality between the teachers using the

different assessment strategies was the same: use the assessment to show what students know. Regardless, the initial comment of a "necessary evil" was a very interesting yet extreme view and could well have reflected how this teacher and their colleagues felt about using the traditional assessment strategy in their classrooms. If teachers really did view the current use of assessments in their classroom as inherently evil, it would likely be seen in how they taught their classes. Thus, it would have an impact on the students whom they taught.

Differences in Who Teaches the Class

Major differences were found between students' perspectives and the teachers whose class they were in for both tests and quizzes (see Table 14 below). There are two hypothetical reasons for these differences, with the first being somewhat logical: the assessment strategies influenced how the teacher taught, thus impacting students' views of their classroom assessments. This suggested the assessment strategy the teacher used in this school influenced not only how students were assessed but also how the teacher taught. Thus, the differences between students' perspectives were rooted in which assessment strategy was being used in the classroom. Table 14 shows positive quotations for Teachers 24 and Diet Coke, both of whom used the Dan Meyer Strategy and more negative quotations for Teacher X and Teacher A, who both used the Traditional Strategy.

Table 14

Differences Between Teachers

Teacher	24	Diet Coke	Teacher A	X
Student claims	"She's a good teacher." "She never gets frustrated with us. That's one of my favorite things."	"I have never felt as confident as I do now because he is my teacher."	"And you know there are people who are scared to stand up and say hey, I don't get this… I am one of them."	"And it's kind of like, if you want me to get an 'A' on this test, you have to give me the tools I can use. And he doesn't."

Note. Teacher and Test Factor Items – $F(3, 405) = 60.21$, $p = .00$
Teacher and Quiz Factor Items – $F(3, 402) = 45.30$, $p = .00$

The second theory of why differences existed between students' views and who taught their math class could be based simply on teaching style: teachers have different personalities. According to Rushton, Morgan, and Richard (2007), "The knowledge and understanding of ones [personality] type is pertinent to success in public education" (p. 440), suggesting teachers' personalities did impact classrooms. It allowed the researcher to gently make the connection that the personality of the teacher could influence how their students viewed their class. The quotations for each educator in Table 14 were how students generally described their teachers. These could simply be a reflection of the personalities that agreed and disagreed with students.

Differences Between Genders

The most perplexing differences that appeared in this doctoral study dealt with the differences in perspectives of male math students and female math students. Not only did differences emerge but they emerged for both the test factor and quiz factor items on the survey (see Table 15).

Table 15

Differences Between Genders

	Test Factor Items Males – M = 32.16, SD = 6.69 Females – M = 33.68, SD = 6.82	Quiz Factor Items Males – M = 28.46, SD = 6.78 Females – M = 30.38, SD = 6.79
Interviewer Questions	When you think of tests, what comes to mind?	When you think of quizzes, what comes to mind?
24	(In same conversation with 2 sentences omitted.) Male: "Not important. I think it's not important." Female: "I guess they can help you learn taking all these tests and quizzes and seeing what you did wrong."	(In exact conversational order.) Male: "It's not important. You might as well just give us the test. They're worth more." Female: "It's kind of like English or Social Studies; they give you a study guide so it's kind of like that. So you can learn and do better on your tests.
Diet Coke	(In same conversation with 3-4 sentences omitted.) Male: "You see one problem four times…" Female: "Because there are just some problems that just show up."	Female: "The other ones [tests] feel more like quizzes since we have them once a week."
Teacher A	(In same conversation.) Female: "He just tells us it's going to be easy and all this kind of stuff and when we actually get to it, no it's not. It's hard." Male: "I don't know. When I took it, it was pretty easy."	(In exact conversational order.) Male: "Before we take the quizzes, we always review." Female: "And then we still fail." Different Male: "I don't."
X	(In exact conversational order.) Female: "I think about cramming things at the last minute and then trying to remember all of it." Male: "I generally don't even think about it because I kind of view it as like, okay, we've already learned this."	(In exact conversational order.) Female 1: ""They're a waste of time." Female 2: "The thing is the tests and quizzes destroy my grade." Male: "I feel like the quizzes though you get them before the test, so if you bomb the quiz you know what to do better on the test."

Note. Gender & Test Factor Items – $F(1, 402) = 5.15$, $p = .02$
Gender & Quiz Factor Items – $F(1, 399) = 8.04$, $p = .01$

Finding differences between genders was rather surprising but should not have been unexpected due to previous research (Adams, Thomas, & King, 2000; Roth & Damico, 1999). After the qualitative analysis, it was suspected that quantitative differences would be found between students' perspectives of the two classroom assessment strategies and students' perspectives of their teachers. However, it was not assumed that differences in perspectives would be found between genders. Once the quantitative analysis was completed and differences were found to be significant for both the test factor items, $F(1, 402) = 5.15$, $p = .02$, and quiz factor items, $F(1, 399) = 8.04$, $p = .01$, it was necessary to re-analyze the qualitative focus groups to see if any differences existed and were initially missed. When looking at the different comments between genders for this section from the qualitative strands, it was easy to see how differences existed between males and females, just as the quantitative strands of data inferred. The qualitative data presented in Table 15 showed there were in fact differences between perspectives of genders upon re-analyzing focus group data for males and females.

One area of warning that should be considered is the qualitative differences found between tests and quizzes for Teacher 24 and Diet Coke as these two teachers used the Dan Meyer assessment strategy. Since these teachers used this strategy, there was no distinguishing between tests and quizzes: only one assessment was used in these classes. The response by the females was summed up best by Diet Coke when she said, "The other ones [tests] feel more like quizzes since we have them once a week" on Table 11. Therefore, these comparisons should be taken with caution.

Overall, it was important to find that differences did exist both qualitatively and quantitatively in this study between male and female math students. However, just

because differences did exist, they did not necessarily support one another. The quantitative data suggested females on average tended to view items with higher frequency than males for both the test factor items and the quiz factor items. Qualitative examples on the table refuted the quantitative findings. For example, Teacher A had female and male students with different perspectives; the male student felt tests were "pretty easy" and the female student felt the tests were "hard." These data did not support the quantitative data that suggested females' perspectives were more favorable than males' perspectives on test factor items. The same conflicting results were found for Teacher A and the quiz factor items. A female student suggested students "still fail" even when reviewing is used before quizzes. A male student quickly remarked, "I don't." Therefore, disparity existing between the data strands and these examples showed they did not necessarily agree. However, both qualitative and quantitative data supported differences existing between genders.

The combined analysis of both qualitative and quantitative strands of data showed better depth and understanding to the complex research question: "Are there differences between secondary math students' perspectives of classroom assessments and their age, gender, and assessment strategy, and how do students' perspectives compare with their secondary math teachers' perspectives of classroom assessment practices?" The quantitative results showed that differences existed between groups of students; however there were not enough data to explain why these differences existed. The qualitative data gave depth and understanding as to why the differences existed and the qualitative strands of data supported and explained what the quantitative data inferred. The qualitative data also revealed two different classroom assessment strategies being taught

in the same building: the Dan Meyer Strategy and the Traditional Strategy. If the qualitative data had not been collected and analyzed, the existence of these two strategies would have remained a mystery. Additionally, there was a much better comparison between students' perspectives of their classroom assessments and the views of the teachers on the assessment strategies used in class. These views had similarities and differences that existed between how teachers viewed their assessment strategies compared to how students viewed the assessment strategies they were given. Knowing these similarities and differences exist is important for future math educators.

Implications

This section identifies the major implications, is designed to enhance the discussion between assessment practices in secondary math classrooms, and shares possible theories as to why students preferred the Dan Meyer assessment strategy, why certain teachers might be favored over others, and gives suggestions regarding the differences found between genders.

Favoring the Dan Meyer Strategy

The effects of this research can only be applied to the students and teachers in this doctoral study and might not extend to findings outside of the sample population due to the lack of using the Dan Meyer assessment strategy. Hypothetically, if this study were to treat the Traditional assessment strategy as a control group and the Dan Meyer assessment strategy as a test group, implications could extend beyond the population of this study. Unfortunately, it is rare to find secondary math teachers who implement the Dan Meyer assessment strategy; until more teachers use it, it would be unrealistic to make conjectures. For the purposes of this study, it is believed there were three primary

reasons students preferred the Dan Meyer assessment strategy over the Traditional strategy: more opportunities to be successful, lower test anxiety, and the difference in grading.

It would seem the students in this study strongly preferred to be assessed in their math classes using the Dan Meyer Strategy. As one female student under the Dan Meyer assessment strategy said, "I like doing the weekly quizzes rather than doing the big unit test." The students preferred to be taught a concept in class and then have multiple opportunities to be successful on that concept. They would rather have shorter and more frequent assessments than one large assessment (Fulkerson & Martin, 1981). The Dan Meyer assessment strategy simply offered these opportunities where the Traditional assessment strategy did not.

Some of the other students suggested, "It's not a lot of pressure" when they did the weekly assessments under the Dan Meyer Strategy. Lack of text anxiety might be the most crucial variable as to why students preferred this method and it stemmed from repeated opportunities to be successful. Fulkerson and Martin (1981) found that more frequent testing helped test anxiety, although it was not by large amounts. The reality is the Dan Meyer assessment strategy likely had lower test anxiety.

How the assessments were graded in the Dan Meyer strategy could also have been a major reason why students preferred this strategy. Research suggested alternative grading techniques normally favored the students (Carey & Carifio, 2012). There could also be a link between the different grading system and test anxiety. Test anxiety has been associated with academic performance in other literature (Chapell et al., 2005, Fincham, 1989, Yildirim, Genctanirim, Yalcin, & Baydan, 2008); student views

supported this. With the repeated attempts provided by the Dan Meyer Strategy, students were more likely to be successful on concepts, which also supported findings from Fulkerson and Martin (1981). For example, one student reinforced this: "It helps them feel better because they get a good grade."

All three of these were suitable reasons for students highly favoring the Dan Meyer assessments strategy in this doctoral study and each one supported previous research. It was clear students preferred the Dan Meyer Strategy but there are numerous reasons this could be occurring. It could be any one of the implied reasons or a combination so it would be hard to determine which one without further research with the sample group in this study.

Preferred Teachers

It was clear students in this study preferred certain teachers over others based on the data. It is possible the differences in students' perspectives emerged from how the teacher treated his or her students. Finding a difference in students' perspectives and who taught them was not a staggering find. However, finding significant differences between teachers in this doctoral study was surprising; it could mean students might have a practical reason to want to be in certain teachers' classes. In this building, it appeared students preferred to be taught under Teacher Diet Coke or Teacher 24. It is likely because these teachers have a more positive teaching style. However, it is important to note that both of these teachers used the Dan Meyer assessment strategy. This could be a coincidence. It was readily evident that students preferred the Dan Meyer assessment strategy but it is unknown if the assessment strategy influenced how teachers taught their students.

Using what was found in this doctoral study, there seemed to be a legitimate connection between students' perspectives and how students were treated by their teachers. For example, the second most mentioned thing by teacher participant 24's students was the teaching style and how positive it was toward the students. Comments like "She never gets frustrated with us," "She's a good teacher," and "I like her a lot" all seemed to be more about how students were treated and had no relation to the Dan Meyer assessment strategy. However, it should not be ruled out as research suggested assessment is teaching (Allen et al., 2009). Teacher Diet Coke also had similar comments from students in a class that supported the teaching style used. Still, many of the comments combined the treatment of students with the assessment strategy used by Diet Coke. One such comment really reflected this: "His teaching method and his assessment method is way better than any teacher I've had yet." Other comments from students from class supported the positive teaching style: "He makes sure we understand it before we move on," "He'll continue to teach you even though we've moved passed," "I just get what he's teaching," and "He makes it a really easy environment to ask questions." Teaching style likely had an influence on students' perspectives but the teaching styles could also have been influenced by the assessment strategies used in classrooms. It is safe to conclude the quantitative data showed students preferred to be taught by 24 or Diet Coke but it was unclear how to determine why students felt this way based on the qualitative data.

Gender Differences

It is uncertain why differences in students' perspectives and gender existed in this study. Research suggested finding differences between genders' perspectives are not new

in education (Adams et al., 2000; Roth & Damico, 1999) and this study supported those previous results. Even more so, Roth and Damico (1999) looked specifically at student perspectives of instruction and differences in gender and found significant differences existed as did this study. However, the reasons why the gender differences existed in this study were unclear. The quantitative data identified significant differences existed with females looking more favorably at assessments than males but the qualitative data did not necessarily support those findings. It was also initially unnoticed that differences existed during the qualitative analysis. Unfortunately, not enough data were available to really understand why these differences existed in this current study.

In conclusion, secondary math students greatly favored the Dan Meyer Strategy of classroom assessments over the Traditional Strategy. Students also favored the two teachers' who used the Dan Meyer Strategy. This could be due to personality or attitude but it could also be tied into how assessments were graded and feedback given; thus, the Dan Meyer Strategy was favored by both male and female students based on data. If the Dan Meyer Strategy is favored this highly over the Traditional Strategy of classroom assessments, perhaps changes are in order to better support what secondary math students identified as helping them more in the classroom. It is suggested that secondary math teachers outside of this study "adopt" the Dan Meyer Strategy and implement it into their classrooms to see if students favor the strategy in their own classrooms and also to see if academic gains could be increased.

Teachers and Assessment

Much of the previous literature revealed shortcomings of teachers not understanding assessment well (Guskey, 2003; Heritage, 2007). This study somewhat

supported this literature. Participants Teacher A and Teacher X used the Traditional strategy of assessment in their rooms; there were hints they did not fully see the importance of their assessments and how they could be used as tools to help students. However, teacher participants 24 and Diet Coke used the Dan Meyer strategy which contradicted the literature, showing not only did they recognize the importance of the assessment to themselves but to their students as well. When teacher participant Diet Coke was asked how he first heard of the Dan Meyer strategy, Diet Coke simply replied:

> I spent a lot of time in my early career looking for online resources and stumbled on Dan's blog. I've read nearly everything he's posted on there at some point. I also saw his TED talk on math classrooms and that sort of hammered it home, for me, that he knows what he's talking about in math instruction. So I adopted his assessment strategy to try it out and never looked back.

Teacher participant Diet Coke went about finding this strategy without a workshop and without training. There are some teachers who can develop and change their practice of assessments and classroom without professional development or professional training.

Limitations and Future Research

There were some limitations of this doctoral study as suggested in the discussions above. Overall, seven limitations to this study were found. The first was the lack of data to fully understand the reasons why students preferred the Dan Meyer assessment strategy over the Traditional Strategy. Understanding the nuances of the Dan Meyer Strategy and what made it so much more attractive to students should be pursued. Another area to explore with more depth is determining why students preferred to be in certain educators' classrooms: was it because of the assessment strategy or was it more how students were treated? Having different questions for the focus group to better understand possible answers would be ideal along with a follow-up focus group with the

same students from classrooms to better understand some of the differences after initial qualitative analysis would help pursue possible explanations. Determining the differences between students' perspectives and gender should also be considered more in-depth now knowing the difference exists. Perhaps having same-gender focus groups form teachers' classrooms would help develop different themes that could be compared to better understand why the differences occurred. Ability level of students was not pursued in this doctoral study due to the disparity of classes taught among teachers. A larger sample of teachers and students could remedy this shortcoming. Similarly, grades and performance scores between students were not considered. Tracking students based on state-wide tests and grades in class might be more intrusive but would yield results that could be analyzed more thoroughly. The survey itself was designed specifically for high school students; it is uncertain if the psychometric results would remain consistent if it was used for middle level students or elementary students. It could also be changed to better reflect the different assessment strategies used in this study.

REFERENCES

Adams, T. L., & Hsu, J. Y. (1998). Classroom assessment: Teachers' conceptions and practices in mathematics. *School Science and Mathematics, 98*(4), 174-180.

Adams, C., Thomas, R., & King, K. (2000). Business students' ranking of reasons for assessment: Gender differences. *Innovations in Education and Training International, 37*(3), 234-243.

Alkharusi, H. (2008). Effects of classroom assessment practices on students' achievement goals. *Educational Assessment, 13*, 243-266.

Alkharusi, H. (2010). Teachers' assessment practices and students' perceptions of the classroom assessment environment. *World Journal on Educational Technology, 2*(1), 27-41.

Alkharusi, H. (2011). Development and datametric properties of a scale measuring students' perceptions of the classroom assessment environment. *International Journal of Instruction, 4*(1), 105-120.

Allen, D., Ort, S. W., & Schmidt, J. (2009). Supporting classroom assessment practice: Lessons from a small high school. *Theory Into Practice, 48*, 72-80.

Ayala, C. C., Shavelson, R. J., Ruiz-Primo, M. A., Brandon, P. R., Yin, Y., & Young, D. B. (2008). From formal embedded assessments to reflective lessons: The development of formative assessment studies. *Applied Measurement in Education, 21*, 315-334.

Bahr, D. L. (2007). Creating mathematics performance assessments that address multiple student levels. *Australian Mathematics Teacher 63*(1), 33-40.

Berenson, S. B., & Carter, G. S. (1995). Changing assessment practices in science and mathematics. *School Science and Mathematics, 95*(4), 182-186.

Berliner, D. (2011). Rational responses to high stakes testing: the case of curriculum narrowing and the harm that follows. *Cambridge Journal of Education, 41*(3), 287-302.

Black, P., & Wiliam, D. (1998a). Assessment and classroom learning. *Assessment in Education, 5*(1), 7-73.

Black, P., & Wiliam, D. (1998b). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappa, 80*(2), 139-144.

Black, P., & Wiliam, D. (2004). The formative purpose: Assessment must first promote learning. *Yearbook of the National Society for the Study of Education, 103*(2), 20-50.

Bol, L., Stephenson, P. L., & O'Connell, A. A. (1998). Influence of experience, grade level, and subject area on teachers' assessment practices. *Journal of Educational Research, 91*(6), 323-330.

Boothroyd, R. A., McMorris, R. F., & Pruzek, R. M. (1992). *What do teachers know about measurement and how did they find out?* Paper presented at the Annual Meeting of the National Council on Measurement in Education, San Francisco, CA.

Britton, T. (2011). Using formative and alternative assessments to support instruction and measure student learning. *Science Scope, 34*(5), 16-21.

Brookhart, S. M. (1997). A theoretical framework for the role of classroom assessment in motivating student effort and achievement. *Applied Measurement in Education, 10*(2), 161-180.

Brookhart, S. M., & DeVoge, J. G. (1999). Testing a theory about the role of classroom assessment in student motivation and achievement. *Applied Measurement in Education, 12*(4), 409-425.

Brookhart, S. M., & Durkin, D. T. (2003). Classroom assessment, student motivation, and achievement in high school social studies classes. *Applied Measurement in Education, 16*(1), 27-54.

Brown, G. T. L., & Hirschfeld, G. H. F. (2007). Students' conceptions of assessment and mathematics: Self-regulation raises achievement. *Australian Journal of Educational & Development Psychology, 7*, 63-74.

Carey, T., & Carifio, J. (2012). The minimum grading controversy: Results of a quantitative study of seven years of grading data from an urban high school. *Educational Researcher, 41*(6), 201-208.

Chapell, M. S., Blanding, B. Z., Silverstein, M. E., Takahashi, M., Newman, B., Gubi, A., & McCann, N. (2005). Test anxiety and academic performance in undergraduate and graduate students. *Journal of Educational Psychology, 97*(2), 268-274.

Chappuis, S., & Stiggins, R. J. (2002). Classroom assessment for learning. *Educational Leadership, 60*(1), 40-43.

Cizek, G. J., Rachor, R. E., & Fitzgerald, S. (1995). *Further investigation of teachers' assessment practices*. Paper presented at the Annual Meeting of the National Council on Measurement in Education, San Francisco, CA.

Clarke, D. J. (1992). Activating assessment alternatives in mathematics. *The Arithmetic Teacher, 39*(6), 24-29.

Crawford, L., Almond, P., Tindal, G., & Hollenbeck, K. (2002). Teacher perspectives on inclusion of students with disabilities in high-stakes assessments. *Special Services in the Schools, 18*(1/2), 95-118.

Creswell, J. W. (2002). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed.). New York: Pearson.

Creswell, J. W. (2007). *Qualitative inquiry and research design: Choosing among five approaches* (2nd ed.). Thousand Oaks, CA: Sage.

Creswell, J. W., & Plano-Clark, V. L. (2011). *Conducting and designing mixed methods research* (2nd ed.). Thousand Oaks, CA: Sage.

Dunn, K. E., & Mulvenon, S. W. (2009). A critical review of research on formative assessment: The limited scientific evidence of the impact of formative assessment in education. *Practical Assessment Research & Evaluation, 14*(7), 1-11.

Edmunds, H. (1999). *The focus group research handbook*. Lincolnwood, IL: NCT Business Books.

Ferguson, P. (2009). Student perception of quality feedback in teacher education. *Assessment & Evaluation in Higher Education, 36*(1), 51-62.

Fincham, F. D. (1989). Learned helplessness, test anxiety, and academic achievement: A longitudinal analysis. *Child Development, 60*(1), 138-145.

Fluckiger, J., Vigil, Y. T., Pasco, R., & Danielson, K. (2010). Formative feedback: Involving students as partners in assessment to enhance learning. *College Teaching, 58*, 136-140.

Frey, B. B., & Schmitt, V. L. (2010). Teachers' classroom assessment practices. *Middle Grades Research Journal, 5*(3), 107-117.

Fulkerson, F. E., & Martin, G. (1981). Effects of exam frequency on student performance, evaluation of instructor, and test anxiety. *Teaching of Psychology, 8*(2), 90-93.

Gall, M. D., Gall, J. P., & Borg, W. R. (2007). *Educational research: An introduction (*8th ed.). New York: Pearson.

Ginsburg, H. P. (2009). The challenge of formative assessment in mathematics education: Children's minds, teachers' minds. *Human Development, 52*, 109-128.

Greenlees, J. (2011). The fantastic four of mathematics assessment items. *Australian Primary Mathematics Classroom, 16*(2), 23-29.

Gresty, K. A., & Edwards-Jones, A. (2012). Experiencing research-informed teaching from the student perspective: insights from developing an undergraduate e-journal. *British Journal of Educational Technology, 43*(1), 153-162.

Guskey, T. R. (2003). How classroom assessments improve learning. *Educational Leadership, 60*(5), 6-11.

Hakel, M. D. (1968). How often is often?. *American Psychologist, 23*(7), 533.

Harlen, W. (2005a). Teachers' summative practices and assessment for learning: Tensions and synergies. *The Curriculum Journal, 16*(2), 207-223.

Harlen, W. (2005b). Trusting teachers' judgment: Research evidence of the reliability and validity of teachers' assessment used for summative purposes. *Research Papers in Education, 20*(3), 245-270.

Hassan, K. E. 2009). Investigating substantive and consequential validity of student ratings of instruction. *Higher Education Research & Development, 28*(3), 319-333.

Heritage, M. (2007). Formative assessment: What do teachers need to know and do? *Phi Delta Kappa, 89*(2), 140-145.

Heritage, M. (2010). *Formative assessment and next-generation assessment systems: Are we losing and opportunity?* Paper presented at the meeting of the Council of Chief State School Officers, Washington, DC.

Holler, E. W., Gareis, C. R., Martin, J., Clouser, A., & Miller, S. (2008). Teacher-made assessments: Getting them right. *Principal Leadership, 9*(1), 60-64.

Hwang, G., & Chang, H. (2011). A formative assessment-based mobile learning approach to improving the learning attitudes and achievements of students. *Computers & Education, 56*, 1023-1031.

Impara, J. C., Plake, B. S., & Fager, J. J. (1993). Teachers' assessment background and attitudes toward testing. *Theory Into Practice, 32*, 113-117.

Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research, 1*(2), 112-133.

Kingston, N., & Nash, B. (2011). Formative assessment: A meta-analysis and a call for research. *Educational Measurement: Issues and Practice, 30*(4), 28-37.

Lalley, J. P., & Gentile, J. R. (2009). Classroom assessment and grading to assure mastery. *Theory Into Practice, 48,* 28-35.

Linn, R. L., & Miller, M. D. (2005). *Measurement and assessment in teaching.* Upper Saddle River, NJ: Pearson.

Maccini, P., & Gagnon, J. C. (2006). Mathematics instructional practices and assessment accommodations by secondary special and general educators. *Exceptional Children, 72*(2), 217-234.

Marzano, R. J. (1994). Lessons from the field about outcome-based performance assessments. *Educational Leadership, 51*(6), 44-50.

McIntosh, M. E. (1997). Formative assessment in mathematics. *The Clearing House, 71*(2), 92-96.

McMorris, R. F., & Boothroyd, R. A. (1993). Tests that teachers build: An analysis of classroom tests in science and mathematics. *Applied Measurement in Education, 6*(4), 321-342.

McTighe, J., & O'Connor, K. (2005). Seven practices for effective learning. *Educational Leadership, 63*(3), 10-17.

Merriam, S. (2009). *Qualitative research: A guide to design and implementation* (2nd ed.). San Francisco, CA: Jossey-Bass.

Meyer, D. (2007, August 31). *The comprehensive math assessment resource* [Web log comment]. Retrieved from http://blog.mrmeyer.com/2007/the-comprehensive-math-assessment-resource/

Muhr, T. (2014). ATLASti (Version 6.2.28). Retrieved from www.atlasti.com.

Newton, P. E. (2007). Clarifying the purposes of educational assessment. *Assessment in Education, 14*(2), 149-170.

Nitko, A. J., & Brookhart, S. M. (2011). *Educational assessment of students* (6th ed.). Boston, MA: Pearson.

No Child Left Behind Act of 2001, Pub. L. No. 107-110, § 115, Stat. 1425 (2002).

Ohlsen, M. T. (2007). Classroom assessment practices of secondary school members of NCTM. *American Secondary Education, 36*(1), 4-14.

Onwuegbuzie, A. J., Witcher, A. E., Collins, K. M. T., Filer, J. D., Wiedmaier, C. D., & Moore, C. W. (2007). Students' perceptions of characteristics of effective college teachers: A validity study of a teaching evaluation form using a mixed-methods analysis. *American Educational Research Journal, 44*(1), 113-160.

Popham, J. W. (1995). *Classroom assessment: What teachers need to know*. Needham Heights, MA: Allyn and Bacon.

Popham, J. W. (2006, September). *Defining and enhancing formative assessment.* Paper presented at the meeting of the Council of Chief State School Officers, Austin, TX.

Rieg, S. A. (2007). Classroom assessment strategies: What do students at-risk and teachers perceive as effective and useful? *Journal of Instructional Psychology, 34*(4), 214-225.

Romagnano, L. (2001). The myth of objectivity in mathematics assessment. *The Mathematics Teacher, 94*(1), 31-37.

Roth, J., & Damico, S. B. (1999). Student perspectives on learning and instruction: Differences by race/ethnicity and gender. *Journal of at-Risk Issues, 6*(1), 32-39.

Rushton, S., Morgan, J., & Richard, M. (2007). Teacher's Myers-Briggs personality profiles: Identifying effective teacher personality traits. *Teaching and Teacher Education: An International Journal of Research and Studies, 23*(4), 432-441.

Schmidt, M. E., & Brosnan, P. A. (1996). Mathematics assessment: Practices and reporting methods. *School Science and Mathematics, 96*(1), 17-20.

Scriven, M. (1979). Interview: Michael Scriven--Viewpoints on education evaluation. *Educational Evaluation and Policy Analysis, 1*, 66-72.

Seymour, M. W., & Chance, S. (2010). Assessment formats: Student preferences and perceptions. *The International Journal of Learning, 17*(10), 137-154.

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

Shepard, L. A. (2005). *The future of assessment: Shaping teaching and learning*. Paper presented at the ETS Invitational Conference, New York, NY.

Simpson, R. H. (1944). *The specific meanings of certain terms indicating differing degrees of frequency*. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/00335634409381009?journalCode=rqjs20#preview

Skwarchuk, S. (2004). Teachers' attitudes toward government-mandated provincial testing in Manitoba. *Alberta Journal of Educational Research, 50*(3), 252-282.

Stiggins, R. J. (2002). Assessment crisis: The absence of assessment for learning. *Phi Delta Kappan, 83*(10), 758-765.

Stiggins, R. J., & Bridgeford, N. J. (1985). The ecology of classroom assessment. *Journal of Educational Measurement, 22*(4), 271-286.

Stiggins, R. J., & Conklin, N. F. (1992). *In teachers' hands: Investigating the practices of classroom assessment.* Albany, NY: SUNY.

Tang, T. L. (1994). *Teaching evaluation in the College of Business: Factors related to the overall teaching effectiveness*. Retrieved from http://files.eric.ed.gov/fulltext/ED374716.pdf

Taras, M. (2005). Assessment--summative and formative--some theoretical reflections. *British Journal of Educational Studies, 53*(4), 466-478.

Taras, M. (2009). Summative assessment: The missing link for formative assessment. *Journal of Further and Higher Education, 33*(1), 57-69.

Taylor, C. S., & Nolan, S. B. (2005). *Classroom assessment: Supporting teaching and learning in real classrooms*. Upper Saddle River, NJ: Pearson.

Tiknaz, Y., & Sutton, A. (2006). Exploring the role of assessment tasks to promote formative assessment in key stage 3 geography: Evidence from twelve teachers. *Assessment in Education, 13*(3), 327-343.

Timmerman, M. E., & Lorenzo-Seva, U. (2011). Dimensionality assessment of ordered polytomous items with parallel analysis. *Psychological Methods, 16*, 209-220.

Van Zoost, S. (2011). Changes and possibilities: A case study of Nova Scotia classroom assessment policies. *Journal of Education Policy, 26*(1), 83-94.

Warren, E., & Nisbet, S. (2001). How grades 1-7 teachers assess mathematics and how they use the assessment data. *School Science and Mathematics, 101*(7), 348-355.

Watson, A. (2006). Some difficulties in informal assessment in mathematics. *Assessment in Education, 13*(3), 289-303.

Watty, K., Jackson, M., & Yu, X. (2010). Students' approaches to assessment in accounting education: The unique student perspective. *Accounting Education: An International Journal, 19*(3), 219-234.

Wiliam, D., Lee, C., Harrison, C., & Black, P. (2004). Teachers developing assessment for learning: Impact on student achievement. *Assessment in Education, 11*(1), 49-65.

Wren, J., Sparrow, H., Northcote, M., &Sharp, S. (2009). Higher education students' perceptions of effective assessment. *The International Journal of Learning, 15*(12), 11-23.

Yang, S. (2010). Teacher evaluation: Teachers' reflection and actions on online student evaluations of teaching. *International Journal of Learning, 17*(1), 133-141.

Yildirim, I., Genctanirim, D., Yalcin, I., & Baydan, Y. (2008). Academic achievement, perfectionism and social support as predictors of test anxiety. *Hacettepe University Journal of Education, 34*, 287-296.

Young, V. M., & Kim, D. H. (2010). Using assessments for instruction improvement: A literature review. *Education Policy and Analysis Archives, 18*(19). Retrieved from http://epaa.asu.edu/ojs.article/view/809

APPENDIX A

GUARDIAN WITHDRAWAL FORM

UNIVERSITY of NORTHERN COLORADO

GUARDIAN WITHDRAWAL FORM

Project Title: Student Perspectives of Assessments
Researcher: Kyle Hemje, College of Education and Behavioral Sciences
Research Advisor: Dr. Harding-DeKam, College of Education and Behavioral Sciences
e-mail: Jennifer.Harding-DeKam@unco.edu

My name is Kyle Hemje. I am a former secondary teacher and currently a graduate student at the University of Northern Colorado. I am working on my Doctorate in Educational Studies and part of my coursework is a research study for my dissertation. I will be coming into your student's classroom and have your child participate in a survey, and then if your child volunteers, to digitally record and lead a small group discussion for practicing purposes.

The purpose of this study is to measure how frequently secondary math students feel the assessments they are given are useful to their learning. Your student's classroom has been selected by approval from both your administration and your teacher who is also participating in the study. The data your student provides will help me gain insight to better understand secondary math students' perspectives. Any information given by your child on the survey cannot and will not be traced back by any parties, including the researcher. I cannot guarantee confidentiality, but in order to protect confidentiality, all information your child provides will be secured and only available to the researcher and the research advisors. All data collected from your student's classroom will be recorded using only numbers, not names. I will not ask for your student's name on the survey at any time or if your child wants to participate in a small group discussion. The only information your student will provide will be his/her gender and age on the survey taken in class. In addition, the use of fake names will be used for the district, school, and teacher to help protect your student.

By giving the survey to your student in the presence of your student's teacher, I will:

- Share who I am and why I am giving the survey.
- Explain how to fill out the survey.
- Distribute the survey to your student.
    - The survey consists of 24 questions regarding how useful your student believes the tests and quizzes are to their learning.
- Collect the survey from your student by passing around a blank folder for him/her to place the survey in.

- Your student's participation is completely voluntary, and no academic penalty will be given if you or your student chooses not to participate.
    - If you or your student chooses not to participate, he/she will remain in the class and simply not fill out the survey. He/she will leave the survey blank so as to protect your student from embarrassment/pressure for not taking it.

Confidentiality will be my top priority throughout the entire study. All data your student provides will be locked in a filing cabinet in my office on the University of Northern Colorado's campus. Data will only be stored in numerical form to protect your student's identity. All electronic data will be secured on the researcher's password protected personal computer.

Your student will be devoting time. Students who choose not to participate should not feel anxious or uncomfortable. However, there is a chance that choosing not to participate while other classmates do may allow your child to feel uncomfortable. To help protect your student from anxiety or discomfort, teachers will be in the classroom during the survey. By choosing to participate, mild anxiety or discomfort associated with taking tests or having a substitute teacher could still occur for your child.

The potential benefits of this study allow great empowerment to your student. These include:

- Student participants' perspectives being used to provide guidance to the clarity and understanding of the survey for future use.
- Giving students opportunities to safely and anonymously identify which assessment practices they feel are most useful to their learning in their classroom.
- In addition, your student will help provide the teacher information which may:
    - Change how the teacher uses assessments in the future.
    - Change the teacher's use of feedback to help students learn.

If either you or your child does not want to participate, please sign the form below. Your child will not be penalized in any way by the school, teacher, or researcher.

Guardian Signature _____ Date _____

APPENDIX B

TEACHER CONSENT FORM

UNIVERSITY of NORTHERN COLORADO

TEACHER PARTICIPATION FORM

Project Title: Student Perspectives of Assessments
Researcher: Kyle Hemje, College of Education and Behavioral Sciences
Phone: (402) 216-7354
e-mail: kyle.hemje@unco.edu
Research Advisor: Dr. Harding-DeKam, College of Education and Behavioral Sciences
e-mail: Jennifer.Harding-DeKam@unco.edu

My name is Kyle Hemje. I am a former secondary teacher and currently a graduate student at the University of Northern Colorado. I am working on my Doctorate in Educational Studies and part of my coursework is to practice my research study that I will use on my dissertation. I am asking for your participation in my research study along with permission to visit your classroom in order for you and your students to participate in a 24 question practice survey. I would also like to have a short conversation with 4 students (2 male and 2 female) who complete the survey from your class. This conversation is to be done right after the students take the survey in your class, and it will be done in your school's library. Therefore, the students will miss a short amount of class time. The conversation will be digitally recorded for the purposes of analyzing and practicing interview questions for future use. In addition, your participation in the study will include completing a 24 question survey and a short interview later in the day at a time that is convenient for you.

The purpose of this study is to determine how frequently secondary students' feel the assessments they are given are useful to their learning, and how the same assessments are considered to be valuable by the teacher. I am asking for permission to give the survey to a class or classes of your choice and for you to complete a survey as well. Any information given by your students' participation or your own participation on the survey cannot and will not be traced back by any parties, including the researcher (myself). I cannot guarantee confidentiality of all participation, but in order to protect the confidentiality of you and your students' responses, all information provided will be secured and only available to the researcher and the research advisors. All survey data will be reported in aggregate form. All discussions and interviews will be digitally recorded for the purposes of better understanding the assessment practices used in your classroom. The use of pseudonyms will be used for all names of participants and locations for the purposed of data reporting.

By agreeing to participate and allowing your students to participate, you will first grant me permission to interact in the following ways:

- I will be allowed to meet on your campus to discuss the research in more detail and to clarify any questions for any administrator, teacher, or participants.
- You will determine a class of students you teach to allow data collection to potentially take place by students.
- I will provide you with the proper number of withdraw forms for students in the class you selected. You will need to hand these out to the students and collect them prior to me giving the survey in your classroom.
- Upon an agreed time with you and at least 48 hours after withdraw forms have been handed out I will give the 24 question survey to the students and a 24 question survey to you.
    - Student participation is completely voluntary, and no academic penalty will be given to students who choose not to participate.
    - The surveys focuses on frequency of:
        - The assessments students take as being useful to their learning
        - The feedback provided on students' assessments as being useful.
        - Your perspectives of assessment practices you give to your students.
- Before giving students the survey, I will collect all withdraw forms that have been completed. I will then lead an introduction (estimated time no more than three minutes) on the survey explaining to students the purpose and rationale.
- Upon completion of the introduction, I will distribute the surveys to the students who have not returned withdraw forms. I will also give you a 24 question survey to complete. Survey completion time by you and your students is estimated to be 5 minutes.
- Once all participants are completed with the survey, I will ask for 4 student volunteers (preferably 2 male and 2 female) who would like to further help me in my study by participating in a short digitally recorded discussion in the school library. This discussion will consist of around 8 questions and will take an estimated 15 minutes to complete. Then I will collect all surveys by passing around a file for students to place their surveys in and answer any questions any participants have. Then the 4 student volunteers will accompany the researcher to the library to conduct a short digitally recorded discussion.
- I will accompany the students back to the classroom upon completion of the discussion.
- I will then leave the campus and meet up with you to have a short semi-structured interview at a later time in the day that is convenient for you.
    - The interview should take an estimated 15 minutes.

Anonymity and confidentiality will be maintained throughout the entire study. All survey data will be reported in aggregate form and locked in a filing cabinet in my office on the University of Northern Colorado's campus. All digitally recorded interviews and discussions will be secured on my personal password protected lap top computer. You will be given a copy of the transcript of the interview I conduct with you for clarity,

appropriateness, and trustworthiness. No other individuals will have access to the recorded conversations other than me and my research advisor.

By granting me permission to do this study in your class or classes, you will be devoting valuable time. You may feel vulnerable to mild anxiety or discomfort during the survey due to the dynamics of allowing the researcher to be leading the introduction with students about the survey and the conducting of the survey. You may also feel anxiety about allowing the researcher to lead a discussion with 4 of your students in the school library. You could feel mild anxiety when completing your own survey. You could potentially feel anxiety when participating in the interview. However, spot checking will occur throughout the interview to make sure you are still comfortable with continuing in the interview. In addition, to help protect you from anxiety or discomfort, you will stay in the classroom the entire time during the survey and while I am in your classroom. The discussion being led by the researcher will be conducted on school grounds (in the library) and there will be numerous students present for the discussion. Pseudonyms will be used for all data reporting purposes.

However, the benefits of this study allow great empowerment to you and your students. These include:

- Student participants' perspectives being used to provide guidance to the clarity and understanding of what assessment practices you use that benefit their learning the most.
- Giving students a safe and unanimous voice indicating which assessment practices they feel are not as useful to their learning.
- You may reflect on your own practices which could lead to:
    - Philosophical changes in how you use your assessments.
    - Changes in how you use feedback on your assessments.

Participation is voluntary. You may decide not to allow your students to participate in this study and if they begin participation you may still decide to stop and withdraw at any time. Your decision and the decision of your students will be respected and will not result in loss of benefits to which you or your students are otherwise entitled. Having read the above and having had an opportunity to ask any questions, please sign below if you are willing to grant permission for me to contact your students to participate in this research. A copy of this form will be given to you to retain for future reference. If you have any concerns about the selection or treatment as research participants, please contact the Office of Sponsored Programs, Kepner Hall, University of Northern Colorado Greeley, CO 80639; 970-351-2161.

Teacher Signature _____ Date _____

Researcher Signature _____ _____ Date _____

APPENDIX C

STUDENT SURVEY

The following survey is for secondary students. The purpose is to determine how frequently students feel the assessments they are given in class are useful to their learning.

Use the following scale for all questions:

1=Rarely 2=Some of the time 3=Frequently 4=Almost Always

Example Question:

I enjoy my time here at school. 1 2 3 (4)

Since 4 was circled it is assumed that you enjoy your time at school almost always.

Tests in this class:	Rarely	Some of the time	Frequently	Almost always
1.) The tests in class help me learn.	1	2	3	4
2.) The tests in class are useful to me.	1	2	3	4
3.) The tests in class help me understand what I have learned.	1	2	3	4
4.) The tests in class help me understand where I make mistakes.	1	2	3	4
5.) My teacher uses tests in class to give me general feedback on my learning.	1	2	3	4
6.) When I get a test back I know where I made my mistake.	1	2	3	4
7.) When I get a test back I know how to do better next time.	1	2	3	4
8.) I get an opportunity to relearn information on a test when it is returned to me.	1	2	3	4
9.) I get an opportunity to retake tests.	1	2	3	4
10.) Tests in this class are a waste of time.	1	2	3	4
11.) I like this teacher.	1	2	3	4

Quizzes in this class:	Rarely	Some of the time	Frequently	Almost always
1.) The quizzes in class help me learn.	1	2	3	4
2.) The quizzes in class are useful to me.	1	2	3	4
3.) The quizzes in class help me understand what I have learned.	1	2	3	4
4.) The quizzes in class help me understand where I make mistakes.	1	2	3	4
5.) My teacher uses quizzes in class to give me general feedback on my learning.	1	2	3	4
6.) When I get a quiz back I know where I made my mistake.	1	2	3	4
7.) When I get a quiz back I know how to do better next time.	1	2	3	4
8.) I get an opportunity to relearn information on a quiz when it is returned to me.	1	2	3	4
9.) I get an opportunity to retake quizzes.	1	2	3	4
10.) Quizzes in this class are a waste of time.	1	2	3	4

Homework in this Class:	Rarely	Some of the time	Frequently	Almost always
1.) The homework in class helps me learn.	1	2	3	4
2.) The homework in class is a waste of time.	1	2	3	4
3.) The homework in class is used to give me feedback on my learning.	1	2	3	4

Demographics

Circle the following: Male Female

Age: _____

Ethnicity: _____

APPENDIX D

TEACHER SURVEY

The following survey is for secondary teachers. The purpose of this survey is to compare the responses with that of a parallel survey given to the students for the purposes of triangulation and clarity.

Use the following scale for all questions:

1=Rarely 2=Some of the time 3=Frequently 4=Almost Always

Example Question:

I enjoy my time teaching. 1 2 3 (4)

Since 4 was circled it is assumed that you enjoy your time teaching almost always.

Tests in this class:	Rarely	Some of the time	Frequently	Almost always
1.) The tests I give help students learn.	1	2	3	4
2.) The tests I give are useful to students.	1	2	3	4
3.) The tests I give in class help students understand what they have learned.	1	2	3	4
4.) The tests I give in class help students understand where they make mistakes.	1	2	3	4
5.) I use tests to give general feedback on what students are learning.	1	2	3	4
6.) When I give a test back students know where they made their mistakes.	1	2	3	4
7.) When I give a test back students know how to do better next time.	1	2	3	4
8.) Students get an opportunity to relearn information on a test when it is returned to them.	1	2	3	4
9.) Students get opportunities to retake tests.	1	2	3	4
10.) Tests in this class are a waste of time for students.	1	2	3	4
11.) I like my students.	1	2	3	4

Quizzes in this class:	Rarely	Some of the time	Frequently	Almost always
1.) The quizzes I give in class help students learn.	1	2	3	4
2.) The quizzes I give in class are useful to students.	1	2	3	4
3.) The quizzes I give in class help students understand what they have learned.	1	2	3	4
4.) The quizzes I give in class help students understand where they make mistakes.	1	2	3	4
5.) I use quizzes to give general feedback on what students are learning.	1	2	3	4
6.) When I give a quiz back students know where they made their mistake.	1	2	3	4
7.) When I give a quiz back students know how to do better next time.	1	2	3	4
8.) Students get an opportunity to relearn information on a quiz when it is returned to them.	1	2	3	4
9.) Students get an opportunity to retake quizzes.	1	2	3	4
10.) Quizzes in this class are a waste of time for students.	1	2	3	4

Homework in this Class:	Rarely	Some of the time	Frequently	Almost always
1.) The homework in class helps my students learn.	1	2	3	4
2.) The homework in class is a waste of students time.	1	2	3	4
3.) The homework in class is used to give students feedback on their learning.	1	2	3	4

Demographics

Circle the following: Male Female

Number of years taught: _____

Please list the math classes you teach and the number of students in each class:

1.) _____ Number of students: _____
2.) _____ Number of students: _____
3.) _____ Number of students: _____
4.) _____ Number of students: _____
5.) _____ Number of students: _____

APPENDIX E

FOCUS GROUP AND TEACHER INTERVIEW QUESTIONS

Focus Group Questions
Semi-Structured Format

Read Aloud: "Thank you all for volunteering to help me with this study. Please know the purpose of this focus group is to better understand your perspectives. So this part of the study is simply a short discussion about the survey you just took. Before we begin, what questions do you have for me?"

1. This first section talks about tests. So when you think of a test in this math class, what comes to mind?
2. What kind of feedback did you get on your last test in this math class? Describe a specific example where you took a test and got it back from your math teacher and explain why was it helpful/harmful?
3. The second section dealt with quizzes. When you think of quizzes in this math class, what comes to mind?
4. What kind of feedback did you get on the last quiz in this math class?
5. How would you describe the homework in this math class?
6. What are some other ways your math teacher might assess your learning that maybe were not covered on the survey?
7. What is most useful to your learning in your math class?

Teacher Interview
Interview Questions
Semi-Structured Format

Read Aloud: "Thank you for giving up your time to help me with this study. Please know that the purpose of this interview is to better understand your assessment practices. Before we begin, what questions do you have for me? Do you consent to the recording of this conversation for the purpose of this research study?"

1. Can you describe some of the assessment practices you use in your classes? How often do you use these assessment practices?
2. Which of your assessment practices do you consider most useful to you? Why?
3. On the last quiz you gave the students, how did you communicate to them?
4. How do your assessment techniques vary between your math classes?
5. Which of your assessment practices do you consider to be most useful to your students? Why?
6. Explain to me how you use homework in your classes.
7. In your opinion, how does your curriculum help or hinder how your use assessments in your classes?
8. Describe how you design your tests.
9. How do you design your quizzes?
10. What is your purpose of giving your last quiz?

APPENDIX F

INSTITUTIONAL REVIEW BOARD APPROVAL AND SCHOOL PERMISSION FORM

UNIVERSITY of NORTHERN COLORADO

Institutional Review Board

DATE:	April 24, 2013
TO:	Kyle Hemje
FROM:	University of Northern Colorado (UNCO) IRB
PROJECT TITLE:	[436886-2] Student Perspectives of Classroom Assessment
SUBMISSION TYPE:	Amendment/Modification
ACTION:	APPROVED
APPROVAL DATE:	April 23, 2013
EXPIRATION DATE:	July 30, 2013
REVIEW TYPE:	Expedited Review

Thank you for your submission of Amendment/Modification materials for this project. The University of Northern Colorado (UNCO) IRB has APPROVED your submission. All research must be conducted in accordance with this approved submission.

This submission has received Expedited Review based on applicable federal regulations.

Please remember that informed consent is a process beginning with a description of the project and insurance of participant understanding. Informed consent must continue throughout the project via a dialogue between the researcher and research participant. Federal regulations require that each participant receives a copy of the consent document.

Please note that any revision to previously approved materials must be approved by this committee prior to initiation. Please use the appropriate revision forms for this procedure.

All UNANTICIPATED PROBLEMS involving risks to subjects or others and SERIOUS and UNEXPECTED adverse events must be reported promptly to this office.

All NON-COMPLIANCE issues or COMPLAINTS regarding this project must be reported promptly to this office.

Based on the risks, this project requires continuing review by this committee on an annual basis. Please use the appropriate forms for this procedure. Your documentation for continuing review must be received with sufficient time for review and continued approval before the expiration date of July 30, 2013.

Please note that all research records must be retained for a minimum of three years after the completion of the project.

If you have any questions, please contact Sherry May at 970-351-1910 or Sherry.May@unco.edu. Please include your project title and reference number in all correspondence with this committee.

UNIVERSITY of NORTHERN COLORADO

Institutional Review Board

DATE:	July 29, 2013
TO:	Kyle Hemje
FROM:	University of Northern Colorado (UNCO) IRB
PROJECT TITLE:	[436886-3] Student Perspectives of Classroom Assessment
SUBMISSION TYPE:	Continuing Review/Progress Report
ACTION:	APPROVED
APPROVAL DATE:	July 26, 2013
EXPIRATION DATE:	July 26, 2014
REVIEW TYPE:	Expedited Review

Thank you for your submission of Continuing Review/Progress Report materials for this project. The University of Northern Colorado (UNCO) IRB has APPROVED your submission. All research must be conducted in accordance with this approved submission.

This submission has received Expedited Review based on applicable federal regulations.

Please remember that informed consent is a process beginning with a description of the project and insurance of participant understanding. Informed consent must continue throughout the project via a dialogue between the researcher and research participant. Federal regulations require that each participant receives a copy of the consent document.

Please note that any revision to previously approved materials must be approved by this committee prior to initiation. Please use the appropriate revision forms for this procedure.

All UNANTICIPATED PROBLEMS involving risks to subjects or others and SERIOUS and UNEXPECTED adverse events must be reported promptly to this office.

All NON-COMPLIANCE issues or COMPLAINTS regarding this project must be reported promptly to this office.

Based on the risks, this project requires continuing review by this committee on an annual basis. Please use the appropriate forms for this procedure. Your documentation for continuing review must be received with sufficient time for review and continued approval before the expiration date of July 26, 2014.

Please note that all research records must be retained for a minimum of three years after the completion of the project.

If you have any questions, please contact Sherry May at 970-351-1910 or Sherry.May@unco.edu. Please include your project title and reference number in all correspondence with this committee.

UNIVERSITY of NORTHERN COLORADO

SCHOOL PERMISSION FORM

Project Title: Student Perspectives of Assessments
Researcher: Kyle Hemje, College of Education and Behavioral Sciences

Phone: (402) 216-7354
e-mail: kyle.hemje@unco.edu
Research Advisor: Dr. Harding-DeKam, College of Education and Behavioral Sciences
e-mail: Jennifer.Harding-DeKam@unco.edu

My name is Kyle Hemje. I am a former secondary teacher and currently a graduate student at the University of Northern Colorado. I am working on my Doctorate in Educational Studies and part of my coursework is to practice my research study that I will use on my dissertation. I am asking for permission to contact the mathematic department teachers in your school to determine if they are willing to participate and to allow their students to participate in this study.

The purpose of this study is to determine how frequently secondary math students feel the assessments they are given are useful to their learning, and how the same assessments are considered to be valuable by the teacher. The data analyses will determine which assessment practices help students learn best. Any information given by participants on the survey cannot and will not be traced back by any parties, including the researcher. Any information given by the focus group of students and the interview with the teacher will not be used by any party other than the researcher. I cannot guarantee confidentiality of all participants' responses, but in order to protect confidentiality, all information participants provide will be secured and only available to the researcher and the research advisor. All survey data will be reported in aggregate form, and all digitally recorded focus group discussion and interviews with the teacher will be secured on the researcher's password protected computer. Also, the use of pseudonyms will be used for all names of participants and locations for data reporting.

By agreeing to allow me to conduct research, you will first be allowing me to interact with your staff. You or your staff may meet with me at any time in the course of the research to clarify or stop the research process. If you approve, the process will continue in the following manner:

- I will be allowed to meet on your campus to discuss the research in more detail and to clarify any questions for any administrator, teacher, or participants.
- The math teachers who are willing to participate and allow my access to their students will determine a class of students they teach to complete a 24 question survey and complete a 24 question survey themselves.
- I will provide withdraw forms for each student participant in the teachers' chosen classes.
- The teacher will distribute and collect the consent/assent forms prior to the researcher coming to the classroom.
- Upon the agreed time with the participating teacher and at least 48 hours after withdraw forms have been handed out I will give the 24 question survey to the students and participating teacher.
    - Student participation is completely voluntary, and no academic penalty will be given to students who choose not to participate.
    - The surveys focuses on frequency of:
        - The assessments students take as being useful to their learning.
        - The feedback provided on students' assessments as being useful.
        - The teacher's perspectives of assessment practices given to students.
- Before giving students the survey, I will collect all withdraw forms that have been completed. I will then lead an introduction (estimated time no more than three minutes) on the survey reinforcing to students the purpose, rationale, and how I would like students to identify any aspects of the instrument which may be unclear or confusing to them.
- Upon completion of the introduction, I will distribute the surveys to the students who have completed the consent/ascent forms. I will also give the teacher a 24 question survey to complete. Survey completion time by the teacher and his/her students is estimated to be 5 minutes.
- Once all participants are completed with the survey, I will ask for 4 student volunteers (preferably 2 male and 2 female) who would like to further help me in my study by participating in a short digitally recorded discussion in the school library. This discussion will consist of around 8 questions and will take an estimated 15 minutes to complete. Then I will collect all surveys by passing around a file for students to place their surveys in and answer any questions any participants have. Then the 4 student volunteers will accompany the researcher to the library to conduct a short digitally recorded discussion.

- I will accompany the students back to the classroom upon completion of the discussion.
- I will then leave the campus and meet up with the participating teacher to have a short semi-structured interview at a later time in the day that is convenient for them.
    - The interview should take an estimated 15 minutes.

Anonymity and confidentiality will be maintained throughout the entire study. All survey data will be locked in a filing cabinet in my office on the University of Northern Colorado's campus. All digitally recorded interviews and discussions will be secured on my personal password protected lap top computer. No other individuals will have access to the recorded conversations other than me and my research advisor.

By giving permission for me to do research, your staff will be devoting valuable time. Teachers and students who participate may feel vulnerable to mild anxiety or discomfort during the survey due to the dynamics of allowing the researcher to be leading the explanation about the survey and distributing the instrument to the students and teacher. Teacher participants may also feel anxiety about allowing the researcher to lead a focus group discussion with 4 volunteer students in the school library. Those same students may feel mild anxiety about having a discussion with the researcher. To help protect teacher and student participants from anxiety or discomfort, teachers will stay in the classroom during the surveys. The focus group with students will be done on school property, and all students will remain together at all times during the discussion. Pseudonyms will be used for all data reporting purposes.

The potential benefits of this study allow both participants and the school to benefit. These include:

- Student participants' perspectives being used to provide guidance to the clarity and understanding what assessment practices teachers use that benefit their learning the most.
- Giving students a safe and unanimous voice to which assessment practices they feel are not as useful to their learning.
- Teachers who allow access to their classroom may reflect on their own practices which could lead to:
    - Philosophical changes in how teachers use their assessments.
    - Changes in how teachers use feedback on their assessments.
- Potential professional development topics for the school including: assessment strategies for how students learn best.
- Identifying which teachers strongly understand assessment practices and conduct assessment practices that work.

- Reflection on curriculum by both teachers and unique student perspectives for greater insight as to what works with assessment practices and what doesn't.

Participation is voluntary. You may decide not to allow your staff or students to participate in this study and if they begin participation you may still decide to stop and withdraw at any time. Your decision, the decision of your staff, and the decision of your students will be respected and will not result in loss of benefits to which you or your students are otherwise entitled. Having read the above and having had an opportunity to ask any questions, please sign below if you are willing to grant permission for me to contact your students to participate in this research. A copy of this form will be given to you to retain for future reference. If you have any concerns about the selection or treatment as research participants, please contact the Office of Sponsored Programs, Kepner Hall, University of Northern Colorado Greeley, CO 80639; 970-351-2161.

Administrator Signature _____ Date _____

Researcher Signature _____ Date _____

APPENDIX G

CONCEPT LIST

Algebra 1B Concept List

	Concept	Level of Mastery					Reference
1	Integer Operations						
2	Plotting Points						
3	Evaluating Expressions						
4	Solving Equations I						
5	Rewrite Formula						
6	Solving Equations II						
7	Proportions						
8	Unit Conversions						
9	Simple Inequalities						
10	Two-sided Inequalities						
11	Compound Inequalities						
12	Functions						
13	Graphing Functions						
14	Graphing Lines						
15	Finding Slope						
16	Slope-Intercepts Lines						
17	Point-Slope Lines						
18	Line of Best Fit						
19	Systems of Equations						
20	Systems of Inequalities						
21	Properties of Exponents						
22	Rational Exponents						
23	Exponential Functions						
24	Polynomial Operations						
25	Factoring Polynomials						
26	Factoring Quadratics (L.C. = 1)						
27	Factoring Quadratics (Any L.C.)						
28	Displaying Data						
29	Probability						
30	Solving Quadratic Equations by Factoring						
31	Completing the Square						
32	The Quadratic Formula						
33	Absolute Value Inequalities						
34	Parallel Lines						
35	Perpendicular Lines						
36	Graphing Quadratics						
37	Arithmetic Sequences						
	Geometric Sequences						
	Transforming Functions						