THE IMPACT OF COMPUTER-ADAPTIVE BENCHMARK DATA AND ASSESSMENT LITERACY ON STUDENT ACHIEVEMENT AND MOTIVATION IN MATHEMATICS

by

Sheryl J. Rushton

A dissertation submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Education

Approved:

James Dorward, Ph.D.  Amy Brown, Ed.D.
Major Professor  Committee Member

Kerry Jordan, Ph.D.  Cathy Maahs-Fladung, Ed.D.
Committee Member  Committee Member

Susan Turner, Ph.D.  Mark R. McLellan, Ph.D.
Committee Member  Vice President for Research and Dean of the School of Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah
2014
ABSTRACT

The Impact of Computer-Adaptive Benchmark Data and Assessment Literacy on Student Achievement and Motivation in Mathematics

by

Sheryl J. Rushton, Doctor of Philosophy
Utah State University, 2014

Major Professor: Dr. James Dorward
Department: Teacher Education and Leadership

Over the past several decades there has been an emphasis in educational research on student assessment and achievement in mathematics. Formative assessments are designed to inform the instructional decision making process and require assessment literacy to interpret and use data provided by these assessments. Many teachers and students were lacking assessment literacy; therefore, they were unable to adjust their instruction and study habits to increase student performance on summative assessments.

This study investigated the impact Scholastic Math Inventory (SMI) benchmarks and assessment literacy training had on summative assessments and student motivation in mathematics. The researcher analyzed unit posttest scores and results from the Instructional Materials Motivation Survey (IMMS) for seventh- and eighth-grade mathematics control and treatment groups. The study took place in a public International Baccalaureate (IB) charter school that served families from suburban communities in
northern Utah.

An analysis of covariance (ANCOVA), using previous CRT math scores as the covariate, was conducted to determine whether there was a difference in the scores of the students who received the SMI benchmarks and assessment literacy training and the students who did not receive this treatment. An analysis of variance (ANOVA) tested whether there was a difference between the IMMS scores for the students who received the SMI benchmarks and assessment literacy training versus students who did not receive this treatment.

The study results indicated that summative scores for seventh- and eighth-grade students who received instruction for the unit along with SMI benchmark and assessment literacy training were not statistically different from students in the control group. The results also showed that the student mathematical motivation overall mean scores were not statistically significant. However, the subscale of satisfaction did show a significant difference in the means. The researcher recommended that use of SMI and assessment literacy training be examined carefully, as these strategies may not improve summative assessment scores in all cases.

(154 pages)
PUBLIC ABSTRACT

The Impact of Computer-Adaptive Benchmark Data and Assessment Literacy on Student Achievement and Motivation in Mathematics

by

Sheryl J. Rushton, Doctor of Philosophy

Utah State University, 2014

Over the past several decades there has been an emphasis in educational research on student assessment and achievement in mathematics. Formative assessments are designed to inform the instructional decision making process and require assessment literacy to interpret and use data provided by these assessments. Many teachers and students were lacking assessment literacy; therefore, they were unable to adjust their instruction and study habits to increase student performance on summative assessments.

This study investigated the impact Scholastic Math Inventory (SMI) benchmarks and assessment literacy training had on summative assessments and student motivation in mathematics. The researcher analyzed unit posttest scores and results from the Instructional Materials Motivation Survey (IMMS) for seventh- and eighth-grade mathematics control and treatment groups. The no-cast study took place in a public International Baccalaureate (IB) charter school which serves families from suburban communities in Northern Utah.

An analysis of covariance (ANCOVA), using previous CRT math scores as the covariate, was conducted to determine whether there was a difference in the scores of the students who received the SMI benchmarks and assessment literacy training and the students who did not receive this treatment. An analysis of variance (ANOVA) tested whether there was a difference between the IMMS scores for the students who received the SMI benchmarks and assessment literacy training versus students who did not receive this treatment.

The study results indicated that summative scores for seventh- and eighth-grade students who received instruction for the unit along with SMI benchmark and assessment literacy training were not statistically different from students in the control group. The results also showed that the student mathematical motivation overall mean scores were not statistically significant. However, the subscale of satisfaction did show a significant
difference in the means. The researcher recommended that use of SMI and assessment literacy training be examined carefully, as these strategies may not improve summative assessment scores in all cases.
ACKNOWLEDGMENTS

I would like to express my deep gratitude to my major professor and mentor, Dr. James Dorward, for his patient guidance, enthusiastic encouragement, and useful critiques of this dissertation. I would also like to thank the rest of my dissertation committee of Dr. Amy Brown, Dr. Kerry Jordan, Dr. Cathy Maahs-Fladung, and Dr. Susan Turner for their support as I moved from an idea to a completed study. I want to extend special thanks to Dr. Patricia Moyer-Packenham for teaching and guiding me through the Math Education and Leadership program. I also thank the faculty and staff at Channing Hall who encouraged me and lifted my spirits as I tried to balance my teaching responsibilities with my doctoral studies for the past 4 years. My grateful thanks are extended to my students who took part in this study and who had to wait extra days for assignments to be returned because I was doing my homework instead of grading theirs.

Beyond the education and professional support I have received, I have had friends and family who gave constant support and encouragement as I endlessly worked on my doctoral degree. I thank my dear parents, Ray and Karen Tidwell, for making this dream of mine a reality, especially my father for instilling in me the love of learning. I thank my children, Jessica, Breanna, Cayla, and McKell, for their patience as I spent time away from them to complete my studies. You will finally be able to find your mother like the little bird in P. D. Eastman’s “Are You My Mother?” Most of all, I am grateful for the loving support and encouragement given to me by my amazing husband, Clay. Without him, I could not have completed my education.

Sheryl J. Rushton
## CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>iii</td>
</tr>
<tr>
<td>PUBLIC ABSTRACT</td>
<td>v</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>x</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xi</td>
</tr>
<tr>
<td><strong>CHAPTER</strong></td>
<td></td>
</tr>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Background</td>
<td>1</td>
</tr>
<tr>
<td>Problem Statement</td>
<td>4</td>
</tr>
<tr>
<td>Purpose of the Study</td>
<td>4</td>
</tr>
<tr>
<td>Research Questions and Hypotheses</td>
<td>4</td>
</tr>
<tr>
<td>Theoretical Framework</td>
<td>5</td>
</tr>
<tr>
<td>Definition of Terms</td>
<td>7</td>
</tr>
<tr>
<td>Summary</td>
<td>11</td>
</tr>
<tr>
<td>II. LITERATURE REVIEW</td>
<td>12</td>
</tr>
<tr>
<td>Inclusion Criteria</td>
<td>12</td>
</tr>
<tr>
<td>Introduction</td>
<td>13</td>
</tr>
<tr>
<td>Assessment Literacy</td>
<td>15</td>
</tr>
<tr>
<td>Assessments</td>
<td>25</td>
</tr>
<tr>
<td>Mathematical Motivation</td>
<td>39</td>
</tr>
<tr>
<td>Summary</td>
<td>42</td>
</tr>
<tr>
<td>III. METHODOLOGY</td>
<td>45</td>
</tr>
<tr>
<td>Research Questions</td>
<td>45</td>
</tr>
<tr>
<td>Design</td>
<td>46</td>
</tr>
<tr>
<td>Participants</td>
<td>46</td>
</tr>
<tr>
<td>Setting</td>
<td>47</td>
</tr>
<tr>
<td>Material</td>
<td>49</td>
</tr>
<tr>
<td>Data Sources/Instruments</td>
<td>49</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Treatment Group Procedure Timeline</td>
<td>62</td>
</tr>
<tr>
<td>2. Control Group Procedure Timeline</td>
<td>63</td>
</tr>
<tr>
<td>3. Ethnicity of Students by Instructional Group</td>
<td>65</td>
</tr>
<tr>
<td>4. Gender of Students by Instructional Group</td>
<td>65</td>
</tr>
<tr>
<td>5. Grade of Students by Instructional Group</td>
<td>66</td>
</tr>
<tr>
<td>6. Descriptive Statistics for Math CRT Percent Scores, the Covariate, by Instructional Group</td>
<td>67</td>
</tr>
<tr>
<td>7. Descriptive Statistics for Unit Posttest Percent Scores, the Dependent Variable, by Instructional Group</td>
<td>68</td>
</tr>
<tr>
<td>8. Shapiro-Wilk Test of Normality for the Posttest of the Control and Treatment Groups</td>
<td>70</td>
</tr>
<tr>
<td>9. Analysis of Covariance for Unit Posttest Scores, With CRT Mathematics Scores as Covariate</td>
<td>71</td>
</tr>
<tr>
<td>10. Adjusted Means and 95% Confidence Interval for Unit Posttest Scores</td>
<td>72</td>
</tr>
<tr>
<td>11. Descriptive Statistics for IMMS Subscales and Overall Score by Instructional Group</td>
<td>73</td>
</tr>
<tr>
<td>12. Shapiro-Wilk Test of Normality for the IMMS Subscales and Overall Score of the Control and Treatment Groups</td>
<td>74</td>
</tr>
<tr>
<td>13. Levene’s Test of Equality of Variances</td>
<td>75</td>
</tr>
<tr>
<td>14. One-Way Analysis of Variance for the Effects of SMI Benchmark Assessments and Assessment Literacy Training on Four Subscales and the Overall Score</td>
<td>76</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Theory on the impact of benchmarks and assessment literacy on student</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>achievement and motivation</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Flowchart of CAT algorithm</td>
<td>33</td>
</tr>
<tr>
<td>3</td>
<td>Posttest percent scores versus CRT mathematics percent scores</td>
<td>69</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

Background

Over the past several decades there has been an emphasis in educational research on student achievement in mathematics. Fueled in large part by the No Child Left Behind Act (NCLB, 2002), states, districts, and schools have worked to address the call for greater accountability of student achievement and academic growth. The focus of their attention has been on methods and strategies used to measure student achievement (Popham, 2004).

Assessments were believed by some to be a panacea for educational improvement, yet others felt assessments threatened the quality of educational practices (Donhost, 2009). Realizing that there are two main goals of assessment may dispel the incongruence in opinions. Assessments for accountability, or summative assessments, are designed to summarize the achievement level of the students. Assessments for student improvement, or formative assessments, are designed to inform the instructional decision making process (Popham, 2010).

State-level accountability assessments, being summative in nature, lack the ability to inform instruction due to the time it takes to receive the results. Yet, the results of the state-level accountability assessments are used to make high-stakes, often life-changing, decisions (Stiggins, 2007). In response to mounting public and political pressures for schools to improve their state-level assessment results, many school faculty and
administrators are exploring the use of formative assessments (Popham, 2008). Formative assessment refers to assessment for learning and “encompassing all those activities undertaken by teachers, and/or by their students, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged” (Black & Wiliam, 1998a, pp. 7-8). Those activities could include, but are not limited to, homework assignments, quizzes, chapter tests, and benchmark assessments. Traditionally, these have been administered using paper-and-pencil, but an increasing trend in the use of technology has made computer-adaptive-tests (CAT) more prevalent in the classroom (Cohen, 2013; Nugent, 2009; Wainer, 2000).

A great deal of research has identified benefits for using formative assessment (Black & Wiliam, 1998a; Black, Harrison, Lee, Marshall, & Wiliam, 2003). Research findings led to developments of strategies designed to maximize student potentials. Benchmark assessments are one of these strategies. The terms interim assessment, quarterly assessment, and progress monitoring are interchangeable with benchmark assessment. Benchmark assessments are assessments administered periodically throughout the school year to evaluate students’ knowledge and skills relative to an explicit set of longer-term learning goals (Herman, Osmundson, & Dietel, 2010). Proper use of data provided by benchmark assessments can help students improve on their summative assessments (Carlson, Borman, & Robinson, 2011; Karpinski, 2010; Keller-Margulis, Shapiro, & Hintze, 2008; Nugent, 2009).

Many schools administer benchmark assessments and become inundated with student data they do not know what to do with or how to interpret. A cure for this
dilemma is to have both teachers and students become assessment literate (Havnes, 2004; Stiggins, 2007). Assessment literacy is having the knowledge needed to understand and use data provided by formative assessments and benchmarks to aid in adjusting teachers’ instruction, students’ study habits, and gain the required knowledge for success on the summative assessment (Popham, 2008, 2011). A teacher might use their assessment literacy skills by taking the results from a quiz, homework assignment, or some other kind of formative assessment to recognize the areas of mathematical knowledge the class is lacking and adjust their teaching to address those deficiencies. Students might read the score and feedback from a formative assessment to recognize the concepts they missed and use that information to adjust what and how they study the concepts (Popham, 2008).

International tests (TIMSS) and other studies have recognized a decline in the mathematics achievement and motivation of middle school students (Beaton et al., 1996; Martin, Herd, Alagaraja, & Shuck, 2012). Declines found at this grade level have spurred researchers to find ways to improve the middle school mathematics achievement and motivation (Martin et al., 2012). Becoming assessment literate may lead students to have higher self-efficacy in mathematics which, in turn, raises their motivation (Keller, 2010). Many studies have been conducted on student motivation, which is pivotal in educators’ endeavors to improve student achievement. There tends to be a positive relationship between mathematical motivation and academic achievement (Keller, 2010; Plenty & Heubeck, 2013). This relationship is discussed further in the literature review. Teachers use an assortment of pedagogical methods to engage and motivate their students in mathematics. Motivating and engaging students can prove to be a difficult task as many
intrinsic and extrinsic elements and factors can effect student motivation (Mueller, Yankelewitz, & Maher, 2011).

**Problem Statement**

Many teachers and students are lacking assessment literacy, therefore, are unable to adjust their instruction and study habits to increase student performance on summative assessments (Popham, 2008). Self-efficacy and motivation is diminished among middle school students who perform poorly on high-stakes tests (Bandura, 1977). Research has yet to determine the influence that a combination of CAT benchmarks and assessment literacy training with middle school students has on summative test scores.

**Purpose of the Study**

The purpose of this study was to determine the impact of Scholastic Math Inventory (SMI) computer adaptive benchmark assessments and assessment literacy training on seventh- and eighth-grade students’ mathematics achievement and motivation.

**Research Questions and Hypotheses**

1. Do formative SMI computer adaptive benchmark assessments, coupled with assessment literacy training for students, influence seventh- and eighth-grade students’ performance on summative assessments in mathematics?
   
   a. Null hypothesis: The use of SMI and assessment literacy training does not make a difference in student performance on summative assessments in
2. Does training in benchmark assessment literacy increase student motivation?
   a. Null hypothesis: There is no difference in the means scores for motivation between the seventh- and eighth-grade students who received training in benchmark assessment literacy and those in the control groups, as measured by the IMMS.

**Theoretical Framework**

This project was based on Paul Black and Dylan Wiliam’s theoretical framework of *Assessment for Learning*. Black and Wiliam were two of the first and most influential researchers of formative assessment (Popham, 2008). It was suggested by Black and Wiliam (2009) that formative assessment is more an incipient theory of pedagogy than evaluation. The development of *Assessment for Learning* theory assimilated theoretical ideas from several different theorists.

One of the theoretical ideas Black and Wiliam drew upon in order to provide theoretical grounding for formative assessments was Ramaprasad’s (1983) three key processes in learning and teaching. The processes are: establishing where the learners are headed (the teacher’s agenda); establishing where the learners are in their learning (internal world of the student); and establishing what needs to be done for the learner to achieve (the intersubjective; Black & Wiliam, 2009). With these three processes in place, Black and Wiliam posited that the responsibility for learning lies equally with the teacher and the student. The teacher is responsible for designing and implementing an effective
learning atmosphere. The student is responsible for the learning that goes on within that environment (Black & Wiliam, 2009; Leahy, Lyon, Thompson, & Wiliam, 2005; Sadler, 1998).

By understanding where responsibilities lie for learning, Black and Wiliam (2009) were able to develop five key strategies that conceptualize the theory of *Assessment for Learning*:

1. clarifying and sharing learning intentions and criteria for success;
2. engineering effective classroom discussions and other learning tasks that elicit evidence of student understanding;
3. providing feedback that moves learners forward;
4. activating students as instructional resources for one another; and
5. activating students as the owners of their own learning. (pp. 4-5)

Black and Wiliam (2009) drew upon other theorists’ ideas as well. Piaget and Vygotsky provided a theoretical basis with a central aim that encourages cognitive growth by creating conflict that pushes the learner to acquire new mental capabilities (Vygotsky, 1978). By following the five key strategies developed by Black and Wiliam, teachers are able to create the intellectual conflict Piaget and Vygotsky deemed necessary for acquiring new understanding.

Another theoretical basis for this research came from the Cognitive Acceleration Program developed by Shayer and Adey (Adey, 2005; Shayer & Adey, 2002). The program offered a set of pedagogical practices that have significantly improved long-term school achievement, by challenging learners to reflect on their own thinking. Teachers and peers help the learners to make unconscious processes explicit and thus make their knowledge more available for future use. A key aspect of the Cognitive Acceleration Program is that students learn through dialogue with others, both with teachers and with
peers. This follows Vygotsky’s (1978) principle that ideas first come through the external social space, and then they become internalized by the individual.

Black and Wiliam’s theory of Assessment for Learning also incorporated Perrenoud’s (1998) ideas that students need to self-regulate their learning by being assessment literate and use feedback from formative assessment to adjust their learning (Black & Wiliam, 2009; Stiggins & Chappuis, 2006). By combining aspects from the theoretical ideas of Perrenoud, Ramaprasad (1983), and Piaget and Vygotsky (Vygotsky, 1978), Black and Wiliam’s developed a theory of Assessment for Learning which is the theoretical framework for this study.

Research suggests that the use of formative assessments positively correlates to success on summative assessments (Carlson et al., 2011; Karpinski, 2010; Keller-Margulis et al., 2008; Nugent, 2009) leading to improved mathematical motivation (Peetsma & Van der Veen, 2013; Weinberg, Basile, & Albright, 2011). Figure 1 shows the conceptual framework of predicting summative assessment achievement and students’ mathematical motivation using benchmark assessments and students’ interpretation of assessment data, informed by assessment literacy.

**Definition of Terms**

The following are terms defined for this study.

*Adequate yearly progress (AYP):* AYP is the measure by which schools, districts, and states are held accountable for student performance under Title I of the NCLB Act of 2001, the current version of the Elementary and Secondary Education Act (EdWeek, 2011).
Figure 1. Theory on the impact of benchmarks and assessment literacy on student achievement and motivation.

**Benchmark assessment:** Benchmark assessment data show areas in which students are struggling and pinpoint the areas in the curriculum where improvements are needed (Goertz, Oláh, & Riggan, 2009).

**Computer-adaptive-tests (CAT):** Computer-adaptive-tests are designed to match the knowledge and ability of a student by adjusting the level of question difficulty, based on the responses delivered by the test taker (Thompson & Weiss, 2011).

**Criterion-referenced test (CRT):** A criterion-referenced test is the summative assessment administered to students at the conclusion of a course. The score yields an estimate of a student's level of knowledge and skills with respect to a well-defined
domain of content (Swaminathan, 1974).

*High-stakes testing:* A high-stakes test is a test with important consequences for the test taker (Glossary of Educational Reform, 2013).

*Individual education plans (IEP):* An IEP is a written plan/program for students with disabilities developed by the school’s special education team with input from the parents, student, and the general education team. It specifies the student’s academic goals and the plan to obtain these goals (Siegel, 2011).

*International baccalaureate (IB):* The IB Program is an internationally-recognized program established in 1968 and noted for its depth and challenging curriculum. It is a high-quality curriculum that aims to encourage critical thinking by the study of traditional disciplines and encourages an international perspective. IB students are also required to engage in community service, individual research, and inquiry into the nature of knowledge (International Baccalaureate [IB], n.d.).

*Learning progression:* A learning progression is a set of building blocks that students need in order to successfully attain a skill or learning outcome (Popham, 2008).

*Mathematical motivation:* Mathematical motivation is that which “explains what goals people choose to pursue and how actively or intently they pursue them” (Keller, 2010, p. 4) in the subject of mathematics. The amount of mathematical motivation a person has is tied to the person’s self-efficacy in the subject. Mathematical motivation guides the person’s self-regulated learning.

*Reliability:* Reliability for an assessment is the extent to which the assessment produces similar results under repeated consistent conditions (Creswell, 2008).
**Self-efficacy**: Self-efficacy relates to a person’s belief of their ability to do the things necessary for success, to plan for that success, and to sustain that plan (Keller, 2010).

**Self-regulated learning**: Self-regulation is an integrated learning process, guided by motivation to learn. It describes an individual who monitors, directs, and regulates their actions toward a goal of self-improvement. A self-regulated student is aware of their academic strengths and weaknesses (Nicol, 2009; Nicol & Macfarlane-Dick, 2006). A self-regulated learner is one who uses assessment literacy skills.

**Summative assessment**: Summative assessment refers to the assessment of learning. It summarizes the development of students’ learning at a particular time, usually at the conclusion of a unit, section, or course (Popham, 2010).

**Title 1 school**: Title I provides federal funding to schools with the intent to help students who are economically disadvantaged or at risk of falling behind academically. In Utah, the percent of low-income students in a given school must equal or exceed district average poverty. For charter schools to be eligible for Title I funds, the school must have at least 10 low-income students (Utah State Office of Education [USOE], 2011).

**Validity**: The validity of an assessment focuses on the match between the content of the test questions and the knowledge or skills that they are intended to measure, and the match between the collection of test questions, what they measure, and the domain of content that the tests are expected to measure (Creswell, 2008).
Summary

This study addressed the problem that many teachers and students lack assessment literacy and are, therefore, unable to adjust their instruction and learning to increase student performance on summative assessments. Formative benchmark assessments, coupled with assessment literacy in the classroom, have potential to increase student achievement on summative assessments by strengthening the teachers’ and students’ understanding of content knowledge gaps (Black et al., 2003). Teachers and students then use that understanding to adjust instruction and study habits to improve performance on summative assessments in accordance with NCLB. The purpose of this study was to determine the influence of benchmark assessments and assessment literacy on middle school students’ mathematics achievement and motivation.

Chapter II examines relevant literature that supports this study. It examines assessment literacy, assessment (both formative and summative), and motivation. Chapter III outlines the procedures and methodology that were used during the research. Chapter IV presents the results from the data analyses. Chapter V reviews the findings of the study and discusses the impact of these findings.
CHAPTER II

LITERATURE REVIEW

Inclusion Criteria

According to Boote and Beile (2005), a sophisticated literature review includes identifying terms used in the literature search, checking the literature for relevance to the topic of research, organizing the literature that has been selected, and finally writing a literature review. This review of literature was conducted using multiple database searches provided by the Merrill Library at Utah State University and other online sources including ProQuest, Ebsco Host, Education Source, Educational Resources Information Center (ERIC), Educational Full Text, Jstore, and Google Scholar. Key words utilized in the searches were: formative assessment, benchmark assessment, middle school, mathematics, assessment literacy, teacher assessment literacy, student assessment literacy, computer-adaptive assessment, Scholastic Math Inventory (SMI), predicting summative assessment achievement, mathematical motivation, self-regulated learning, and mathematical self-efficacy.

As the initial literature was reviewed, citations and references often led to additional resources that were relevant to the study. Some resources from the initial search were deemed to be irrelevant to the questions addressed in this study. The whole set of literature was then narrowed to the literature most relevant to the topics of interest. At the conclusion of gathering data, another search for current related literature was performed utilizing the same keys words as at the initial literature review.
Introduction

In 2001, the legislation known as NCLB was enacted to increase accountability for school districts across the nation, resulting in federally mandated high-stakes testing in reading, science, and mathematics (NCLB, 2002). High-stakes tests may be used to determine what class a student can enroll in or if a student is allowed to graduate. High-stakes tests are also used to determine if schools have met the required adequate yearly progress (AYP) goals as required by the NCLB act.

Raising the standards of learning and achievement is a national priority. National, state and district standards are set to increase the rigor of the courses. Programs for external testing of students’ performances are being enhanced (Kaufman, Guerra, & Platt, 2006). Over the years, the assessment community has focused on maximizing the efficiency and accuracy of high-stakes tests. Yet, little attention is paid to assessment as it affects teachers and students in daily classroom use (Stiggins, 2007).

Evidence shows that everyday practice of assessment use in classrooms is ridden with problems. These problems include lack of time (Supovitz & Klein, 2003; Wayman & Stringfield, 2006), lack of a technological infrastructure (Chen, Heritage, & Lee, 2005; Lachat & Smith, 2005), and teaching practices that work against use of assessment evidence in an ongoing manner (Ingram, Seashore Louis, & Schroeder, 2004; Supovitz & Klein, 2003; Young, 2006).

For years, an emphasis in educational research has been that of student achievement in mathematics. Assessment is the main area of focus when it comes to measuring student mathematical achievement (Wiliam, Lee, Harrison, & Black, 2004).
Not everyone agrees on the tools needed to increase mathematical achievement (Bernhardt, 2006). Both formative and summative assessments have been under the microscope. Formative assessments are designed to inform instruction and study methods to increase performance levels on summative assessments (Popham, 2010). Traditionally, paper-and-pencil versions of formative assessments (i.e. homework assignments, quizzes, chapter tests, and benchmark assessments) have been used to gather data. Recently, CATs have been added to the list of ways of gathering data from formative assessments. However, many teachers and students do not know what to do with the data obtained from formative assessment (Black, 1993; Popham, 2010, 2011). Students’ achievement on summative assessments often depends on teachers and students being assessment literate. Assessment literacy refers to the understanding teachers and students have to use data provided by formative assessments, such as benchmarks, to adjust learning experiences in order to gain the required knowledge for success on the summative assessment (Popham, 2008, 2011).

Benchmark assessment and other formative assessment data identify areas in which students are struggling and pinpoint the areas in the curriculum where improvements are needed (Goertz et al., 2009). During the past several decades, educators began to discuss the benefits of formative assessment on teachers’ instructional decisions and students’ studying practices (Popham, 2008). Formative assessment is now an integral part of teaching and learning; however, assessment literacy in not yet an integral part of education.

This review examines the literature pertaining to assessment literacy for teachers
and students. The review will then move into a discussion on assessments. The assessment section contains literature on benchmark assessments and assessment with technology, specifically computer adaptive tests. It also contains literature on summative assessments. Research about motivation gained through improved student achievement will then be reviewed.

**Assessment Literacy**

In order to make assessments worthwhile, Popham (2011) posited that teachers and students needed to become assessment literate. “Assessment literacy consists of an individual’s understandings of the fundamental assessment concepts and procedures deemed likely to influence educational decisions” (Popham, 2011, p. 267). According to Havnes (2004), teachers often assumed that it is their teaching that guides the students’ learning. However, in practice it is assessment that directs the students’ learning and defines what is worth learning (Brown, McInerney, & Liem, 2009; Havnes, 2004).

Havnes (2004) conducted an ethnographic study of a compulsory preparatory course at the University of Oslo. He observed, interviewed, and worked side by side with several students in the course. His main argument was that the assessment structure contributed to the establishment of the learning content, how the teachers taught the course, and the students’ learning practices. “Learning is relational. It is relational to assessment, but assessment, again, is relational to other components on the complex system of educational programmes” (Havnes, 2004, p. 171).

In a qualitative study conducted by Lukin, Bandalos, Eckhout, and Mickelson
(2004), a group of teachers were trained in assessment literacy through a formal course offered by the University of Nebraska-Lincoln called Nebraska Assessment Cohort, which was an adaption of the assessment literacy training program developed by Stiggins in 2001. The teachers implemented the skills they learned through the course into their classrooms. At the end of one school year, data was collected from one high school to determine the effectiveness of the training in assessment literacy. The researchers used a questionnaire (Classroom Assessment Questionnaire) and a survey (Self-Assessment Development Levels based on Classroom Assessment Quality Rubrics), both developed by Arter and Busick in 2001. Participants answered several open-ended questions about their skill, confidence levels, and changes they had made in their own classroom teaching and assessment practices. The data collected suggested that the assessment literacy learning training had a positive impact on teacher confidence, knowledge, and skill in the area of classroom assessment. There appears to be evidence, while limited, which suggests students also experienced positive outcomes in achievement.

When teachers use data from an assessment, they can better assist students in their learning progressions (Popham, 2008; Shapiro & Gebhardt, 2012). Data from assessments lead the teacher to develop instruction that is suited to the students’ needs. “Assessment-literate educators come to any assessment knowing what they are assessing, why they are doing so, how best to assess the achievement of interest, how to generate sound samples of performance, what can go wrong, and how to prevent these problems before they occur” (Stiggins, 1995, p. 240). It seems likely that the most self-regulating students use formative assessment to improve the quality of their learning progression
Assessment literacy for teachers is just as important as assessment literacy for students (Popham, 2008; Stiggins, 1995; see Figure 1). Teachers must be assessment literate (Popham, 2011). According to Popham (2011), teachers need to know about the range of assessment strategies so they can maximize the opportunities for gathering evidence. They need to know how to align assessment with instructional goals, and then ensure that inferences drawn from the assessments are of value in aiding the teachers’ understanding of where students are with respect to their learning progressions (Heritage, 2007).

Research suggested that teachers need training to interpret data from assessments, and then take that information to adjust instruction to meet the needs of the students (Blink, 2007; Popham, 2008). Observations and interviews to determine that teachers use the knowledge gained from assessment data to inform their instruction have been used to determine the level of assessment literacy (Christman et al., 2009; Ingram et al., 2004; Shepard, Davidson, & Bowman, 2011). One study (DeLuca & Klinger, 2010) used a questionnaire developed by the researchers to determine the level of assessment literacy for all subject areas in pre-service teachers. The questionnaire focused on the confidence level of differing aspects of several types of assessments. While the pre-service teachers may not have experience in the classroom, they are able to identify their perceived needs, level of knowledge, and sense of readiness pertaining to assessment literacy. The
researchers found that preservice teachers were more confident in their use of assessment of learning (summative) than they were with assessment for learning (formative). The results supported the need for training in assessment literacy.

In a longitudinal study of nine high schools (Ingram et al., 2004), teachers were expected to use data to assess their own, their colleagues', and their schools' effectiveness in all subject areas and to make improvements. The findings suggested that teachers were willing to use the data to make improvements, but they had significant concerns about the kind of information that was available, how it was to be used, and how it would affect their teaching once they had the information provided by the assessment. These findings are consistent with the characteristics that define assessment literacy. The researcher found that when teachers are trained and supported in becoming assessment literate, greater support for students occurred.

Another study was conducted where 24 teachers in six different districts were supported over a 6-month period in exploring and planning their approach to assessments, and learning how to interpret the data (Wiliam et al., 2004). The researchers worked in a collaborative manner with the group of teachers suggesting direction that might be useful to explore. The teachers had 6.5 days of inservice sessions over a year and a half period of time. They developed action plans for implementing formative assessment. Their plans included ideas such as: focusing on or improving the teacher’s own questioning techniques by using more open-ended questions; allowing students more time to think of answers; starting lessons with a focal question; and using comment only marking on assignments and assessments. The teachers put these plans into action with
selected mathematics and science courses. In order to compute effect sizes, at least one comparison group was established for each class. Creating the comparison groups proved to be a challenge and made it difficult to measure effects across all the groups. The researchers had to use the same teacher teaching the same course but in different years, different teachers teaching the same course at parallel times, and the same teacher teaching the same course at parallel times. The results showed “that improving formative assessment does produce tangible benefits in terms of externally-mandated (summative) assessments” (Wiliam et al., 2004, p. 7).

**Teacher implementation of assessment.** The next three studies show how teachers used their assessment literacy to interpret data and become aware of their students’ needs. The teachers then took the knowledge gained through data to inform and adjust instruction to meet those needs.

Wayman and Stringfield (2006) conducted a qualitative study that collected data through focus groups and interviews in three different schools; a pre-kindergarten through grade five elementary school, a large school serving grades five and six, and a middle school grades six though eight. The researchers explored two questions; (a) what facilitates the widespread use of examination and learning from student data, and (b) what changes in faculty practice and attitudes resulted from examining and learning from student data. The researchers found that administration support in fostering examination of student assessment data is vital to widespread use of the data. They also found that data use often resulted in teachers sharing their insights with other teachers which increased their collaboration skills. This, in turn, improved the teachers’ ability to better
understand their students’ abilities. The researchers discovered teachers felt that, along with administrative support, they needed time to learn how to interpret and examine student data. The results indicated teachers were able to use data to go remarkably deep in their examinations of student learning and in their teaching practices. Teacher efficiency was noticeably amplified.

Christman and colleagues (2009) conducted a similar large scale study. They focused on the Philadelphia school district’s use of assessments in three key areas: (a) teachers’ perception of the assessments; (b) how teachers used the assessments; and (c) how the emphasis on data-driven teaching affected the effectiveness of the exams. Their study utilized multimethods that relied on three sources of data: (a) student achievement and demographic data from 2005-2007, (b) district-wide responses to a teacher survey, and (c) qualitative research from 10 schools during the years 2005-2007. The researchers found that benchmark assessments aligned with core curriculum offered the opportunity for teachers to delve deeper into the curriculum as they reviewed the assessment data and became aware of their students’ needs. The most important finding from this study was that the success of benchmark assessments depends on the knowledge and skills of the teachers. Knowledge and skills of the teachers were determined by the amount of training given to the teachers and evidenced in the students’ achievement growth. Christman and colleagues conjectured that “data can make problems more visible, but only people can solve them” (p. 65).

Hoover (2009) used a survey to determine how 1,500 elementary, middle school, and high school teachers in a large suburban school division in central Virginia used
assessments formatively in their classrooms. An overwhelming majority used assessment data to inform and evaluate their instructional practice and make changes to enhance student learning. The results of the study suggested teachers administer formative assessments at varying frequencies throughout the year and analyze the data on a regular basis. The majority of the teachers surveyed reported using central tendency statistics. The Teacher Assessment Literacy Questionnaire, developed by Plake and Impara in 1993, was the instrument used to determine teacher assessment literacy. The researchers found that high school teachers had higher assessment literacy scores than elementary teachers, and teachers with graduate degrees scored higher than those with bachelor’s degrees. The more experienced teachers had higher assessment literacy scores than novice teachers. Finally, the researchers found that mathematics and science teachers had higher assessment literacy scores than other teachers. The teachers with the higher assessment literacy scores implemented formative assessment skills in their classrooms more often and more consistently than the teachers with lower scores.

**Assessment literacy in mathematics.** Assessment literacy supports mathematics teachers in planning their instruction. As Oláh, Lawrence, and Riggan (2010) discovered, teachers analyzed and used data in two ways: (a) to detect errors, concentrating on whether students got problems correct; and (b) to diagnose those errors, focusing on why students might have gotten certain problems wrong. Some teachers looked at the procedures used by students in solving problems, while others focused on underlying mathematical thinking and misconceptions. They found that teachers used their assessment literacy to interpret data from a variety of sources. For example, some
teachers reported asking students to explain responses to particular assessment problems, or encouraging students to show their work. Oláh and colleagues also found that teachers’ analysis of data led to different types of instructional planning.

Shepard and colleagues (2011) observed and interviewed 30 middle school mathematics teachers in seven districts. They discovered that the amount of assessment literacy possessed by teachers determined the extent to which they were able to use data collected by benchmark assessments. Teachers’ uses of assessment information varied; most frequently they retaught standards or items with the lowest scores. Although many teachers expressed an interest in using assessment results to inform instruction, they reported a minimal amount of professional development in assessment literacy. According to Oláh and colleagues (2010) and Shepard and colleagues (2011) and as seen in Figure 1, assessment literacy should inform teachers how to interpret, analyze, and use data from formative assessments to adjust instruction.

**Assessment Literacy for Students**

Not only do teachers need to learn how to use data from assessments, but students also need to learn how to use that information to enhance their achievement (Gibbs & Simpson, 2004; Mac Iver, 1987). Sadler (1998) stated, “Students should be trained in how to interpret feedback, how to make connections between the feedback and the characteristics of the work they produce, and how they can improve their work in the future” (p. 78). Popham’s (2008) statement concurs with Sadler that students must begin the process of using assessment data to improve their work by having a “full-scale orientation” (p. 73) on these learning tactics.
Researchers (Hattie, Biggs, & Purdie, 1996; Heritage, 2007) suggested that the most powerful single moderator to improve achievement is the feedback students get from assessments. It is not common to have students focus on some form of self-assessment or feedback from assessments (Black & Wiliam, 1998b). Yet, Black and Wiliam (1998a) found in their review of literature a large and consistent positive effect on learning from assessment feedback. This conclusion is echoed in many other researchers’ writings since Black and Wiliam conducted their literature review (Black et al., 2003; Popham, 2008; Stiggins, 2007).

Participating in assessments without a perceived purpose, combined with a teacher centered approach to instruction, discourages students from fully engaging in their learning (Robinson & Udall, 2006). Although teachers play an important part in educating students, it is a supporting role (Popham, 2008). When students are trained in assessment literacy, adjustments to learning tactics become student-determined instead of teacher-directed (Brown et al., 2009; Popham, 2008). For many students, assessment is not an educational experience, it is a process of “guessing what the teacher wants” (McLaughlin & Simpson, 2004, p. 136). Robinson and Udall found that students are able to take responsibility for what and how they learn when equipped with the skills to monitor, make judgments, and critically reflect on their performance. These skills include understanding the meaning of the results and feedback from formative assessments and knowing where to look for assistance to fill knowledge gaps when they are discovered in the critical reflection of the results and feedback.

Research has found that providing training to students in assessment literacy can
be beneficial (Brookhart, 2001; McDonald & Boud, 2003; Nicol, 2009; Smith, Worsfold, Davies, Fisher, & McPhail, 2013). Studies have been conducted with high school students (Brookhart, 2001; McDonald & Boud, 2003) while others have used college students as participants (Nicol, 2009; Smith et al., 2013). These studies suggested that assessment literacy is useful and the findings will most likely transfer to other age groups. However, “Neither educational researchers nor educational practitioners fully understand how students’ thoughts, feeling, and actions ultimately influence their academic success” (Artino & Jones, 2012, p. 174).

Artino and Jones (2012), found that boredom, frustration, and low task value interfere with assessment literacy and can be extremely damaging in the learning context. They surveyed 302 undergraduate U.S. service academy students. Even though these students were skilled at assessment literacy, negative emotions made them less likely to employ adaptive learning strategies.

Brookhart (2001) and McDonald and Boud (2003) considered the impact of training high school students on their performance in assessments. Brookhart conducted a qualitative study by interviewing 50 high school students about specific classroom assessment events. The successful students engaged in self-assessment as a regular ongoing process. They studied for tests, they accepted the challenge of mastering difficult material, and they learned on their own by reading resources. The successful students considered these self-assessments or self-regulations as instances of learning.

McDonald and Boud (2003) directed a quasiexperimental study by training 256 high school students in self-assessment skills. The training in McDonald and Boud’s
study focused on similar skills as this current study attempted. The focus was on constructing, validating, applying, and evaluating criteria to apply to students’ work. The researchers surveyed the students to discover their reactions to the training. The survey revealed that training in assessment literacy was a benefit to the students who received it. Both studies, Brookhart and McDonald and Boud, concluded that students having the ability to self-assess and adjust their planning and study habits were more successful in their careers as students. They were able to plan ahead and prepare adequately for exams.

Nicol (2009) and Smith and colleagues (2013) explored how formative assessment and feedback enabled college students to develop their ability for self-regulated learning. Nicol suggested that assessment literacy helps to develop the skills students need to monitor, judge, and manage their own learning. Smith and colleagues’ quasiexperimental study showed how assessment literacy in students contributed to educational gains. The students who received the intervention in assessment literacy were able to develop ability to judge their own and others’ work, which enhanced their learning outcomes.

Assessments

Assessment falls into two different categories: formative and summative. The focus in the next sections will be on these two types of assessments. Formative and summative purposes are different, and thus are usually discussed as two different things (Black, 1998). Formative assessment is designed to provide feedback and to guide in making adjustments in the learning process, both for teachers and for students (Popham,
2008; Stiggins, 2007). Summative assessments measure what was learned after any formative adjustments have been made. “Assessment (formative and summative) is integral to the learning process and something that students ‘take part’ in rather than something that is ‘done to them’” (Robinson & Udall, 2006, p.98).

**Formative Assessments**

The conceptual framework from which the current analysis of literature derives locates formative assessment within the instructional improvement cycle (Figure 1). Many different forms and types of formative assessments exist; however, this study will focus on benchmark assessments administered to middle school students through technology, specifically computer adaptive tests.

Many prominent researchers in assessment (Brookhart, 2001; Gibbs & Simpson, 2004; Heritage, 2007; Kaufman et al., 2006; McDonald & Boud, 2003; Oláh et al., 2010; Popham, 2008; Sadler, 2010; Stiggins, 2007; Stiggins & Chappuis, 2006) credit Paul Black and Dylan Wiliam for piquing the current worldwide interest in formative assessment. Black and Wiliam (1998a) defined “formative” as “encompassing all those activities undertaken by teachers, and/or by their students, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged” (pp. 7-8). Formative assessment is the process of using data about students’ learning to assist teachers to make day-to-day instructional decisions (Black, 1993; Black et al., 2003; Black & Wiliam, 1998a; Heritage & Niemi, 2006; National Research Council, 2001).

A learning environment with formative assessment has many benefits to the
student. In terms of mathematics, research has found that even when the teacher’s mathematical knowledge was low, the use of formative assessments had an underlying capacity to make sense of students’ mathematical understanding and to aid the teacher in responding with the appropriate instruction (Goertz et al., 2009; Hoover, 2009). Popham (2008) boldly stated, “Formative assessment’s raison d’être is to improve students’ learning” (p. 7).

Many researchers have found that formative assessment data were beneficial in several different levels of learning including college, middle school, and elementary (Diefes-Dux, Zawojewski, Hjalmars, & Cardella, 2012; Harlen & Crick, 2003; Koellner, Colsman, & Risley, 2011; Lachat & Smith, 2005; Shepard et al., 2011). Research found feedback from the teachers to be useful to the college students. It helped the students to use their assessment literacy skills when they understood what had been done correctly and what had been done incorrectly from formative assessment feedback (Diefes-Dux et al., 2012). Shepard and colleagues discovered that formative assessments informed teachers of the concepts that needed to be retaught. However, the teachers wanted more professional development to better train them in assessment literacy in order to know how to use the data from formative feedback. Koellner and colleagues conducted case studies and found that when teachers use data from the assessments to identify deficient areas in their students’ content knowledge; they were able to determine instructional methods that could be used effectively with their students.

**Benchmark assessment.** Assessments used for frequent summative “assessments of learning” (Stiggins & Chappuis, 2006, p. 10) may help to evaluate programs, but they
do not help the classroom teacher and student with the current needs of the classroom. However, Stiggins and Chappuis argued that Black and Wiliam’s (1998a) theory of frequent formative “assessments for learning” (p. 11) during instruction could lead to student knowledge gain. Therefore, if benchmarks are used formatively they could improve student achievement.

Formative assessment, in its many forms, has been touted by many as the best way to increase student motivation and achievement in school, particularly in mathematics (Harlen & Crick, 2003; Tonidandel, Quiñones, & Adams, 2002). Several studies examining the relationship between formative benchmark assessments and summative assessments have found positive correlations (Karpinski, 2010; Keller-Margulis et al., 2008; Nugent, 2009). They each were looking for a predictive correlation in order to learn how to identify at risk students.

More recently, Carlson and colleagues (2011), analyzed mathematics summative achievement outcomes using experimental methods. Carlson and colleagues examined mathematical achievement from a district-level random assignment study in over 500 schools within 59 school districts and seven states. Johns Hopkins Center for Data-Driven Reform in Education (CDDRE) provided all the training. The CDDRE instructed the participating districts over a three year period. The first year the districts implemented benchmark assessments and received extensive training on interpreting the data based on the Data-Driven-Design (3D) model, which was created by CDDRE. The second year the districts were to seek out evidenced-based reforms that would address the needs and problems identified by data from the benchmark assessments. The third year the districts
adopted and implemented the program backed by evidence from benchmark data. The study provided evidence that data-driven formative benchmark assessment reforms can result in substantively and statistically significant improvements in summative achievement outcomes.

Research supports the claim that benchmark assessments are useful in assisting teachers design beneficial instruction for their students. Black and Wiliam (2009) noted that teachers may use benchmark assessment to, “make decisions about their next steps in instruction that are likely to be better, or better founded, than the decisions they would have taken in the absence of the evidence” (p. 29). Riggan and Oláh (2011) conducted research where they interviewed and observed 32 teachers in two school districts. They found the most common uses of benchmark assessment data were for determining what to teach and to whom. The teachers used data to identify areas in which students demonstrated weaknesses. Nearly all the teachers reviewed benchmark assessment results by the mathematical content area and by the needs of the student individually. The researchers noticed that externally designed benchmark assessments appeared limited in capacity to inform teachers about students’ thinking or problem solving, but they gave direction to teachers.

Another study by Goertz and colleagues (2009) utilized an in-depth case study of nine elementary schools in two different school districts to address four questions. Two of their questions are relevant to this study. One question examined how teachers, individually and collectively, learn from benchmark assessment results in mathematics and apply that knowledge to instructional decisions about the lesson content, pedagogy,
and how to work with individual students. The second question focused on ways benchmark assessments are situated within the wider context of teachers’ formative assessment practices and tools.

The results indicated that with support through professional development, teachers were attempting to understand data gathered from benchmark assessments, which for some teachers was a new practice. They collaborated with other teachers to increase their instruction building techniques. Goertz and colleagues (2009) noticed that teachers were able to use data from the benchmark assessments to create instruction that connected students’ prior knowledge to current learning goals.

**Assessments with technology.** Lower costs have made the accessibility of computers and other technology affordable to schools (Nugent, 2009; Wainer, 2000). Technology for assessments can include, but are not limited to: electronic portfolios; online assessments (such as the state standardized tests); educational games that track progress; and computer-adaptive tests, which will be discussed in greater detail in a later section.

Effective use of assessment data assumes that the assessments generate useful information on the students’ understanding of the content and information to customize instruction; that teachers and students know how to interpret and act on the assessment data; that results from assessment be received in a timely manner; and that technology will help to streamline these processes (Goertz et al., 2009; Sadler, 2010). Timely, constructive feedback on assessment is an integral aspect of effective and efficient teaching and learning, yet it is difficult to obtain (Black & Wiliam, 1998a, 1998b; Debuse
Research has found that technology helped improve the quantity and timeliness of feedback from formative assessments in the field of mathematics (Bull & McKenna, 2003; Debuse & Lawley, 2011; Denton, Madden, Roberts, & Rowe, 2008; Hepplestone, Holden, Irwin, Parkin, & Thorpe, 2011; Wayman & Stringfield, 2006).

Technology may help students to become better self-regulated learners. Hepplestone and colleagues (2011) posited that technology assessments have the potential to enhance students’ engagement with feedback. The purpose of technology assessments is to encourage independent learning and enable students to monitor their progress in mathematics by taking assessment tasks at a time and place of their choosing (Irons, 2007). The use of computer-assisted assessments allows students to learn and practice the weak areas in their mathematics without the fear of exposing their mistakes to peers (Challis, 2005). The designs of formative computer-assisted assessments could stimulate student commitment to self-evaluation and enhance learning effectiveness (Chen et al., 2005; Wang, Wang, Wang, & Huang, 2006).

Bull and McKenna (2003) composed a list of advantages and disadvantages for the use of computer-assisted assessments. Advantages of assessing with technology include: the promotion of self-monitoring and assessing; detailed feedback is available to the students during the test and immediately after testing; students acquire information technology (IT) skills; large groups can easily be assessed in a quick and consistent manner; and diagnostic reports and analyses can be generated. On the other hand, the disadvantages and limitations of assessing with technology are comprised of: cost and
time-consuming implementation of technology; students must acquire adequate IT skills and experience of the assessment type in advance of the examination; and assessors need training in assessment design, IT skills, and management.

Koedinger, McLaughlin, and Heffernan (2010) conducted a quasiexperimental study where data were collected from 1,240 seventh-grade students in three treatment schools and one comparison school. The study evaluated whether a technology-assisted assessment program, called ASSISTments, had an effect on improving middle school students’ year-end assessment scores. The data determined that students in the treatment schools performed significantly higher than the students in the comparison school. Koedinger and colleagues also discovered that teachers adapted their whole-class instruction based on the data found in the ASSISTments reports. Teachers retaught concepts that normally would not have been revisited. They found teaching techniques that would spark the interest of many of their students. ASSISTment provided teachers with useful formative assessment information. Students who did not use ASSISTments or had low usage benefited from what their teachers learned from observing the students using ASSISTment system regularly and from inspecting the diagnostic reports provided in the data. This study did not address assessment literacy for teachers or students’ motivation toward mathematics.

*Computer-adaptive-tests.* CATs are a type of technological formative assessment that is emerging as an alternative to paper and pencil examinations. Adaptive tests provide the student different test questions according to their performance. They adapt to the level of user abilities and skills. The CAT will randomly present the student with one
or two questions. Based on the answers, the computer will generate the best estimate of
the student’s ability. After each question, the CAT will reassess the student’s ability and
skill level, adapting as needed (Challis, 2005; Hepplestone et al., 2011; Irons, 2007;
Thompson & Weiss, 2011; Tonidandel et al., 2002). The process continues until the
assessment criterion has been met. Figure 2 shows the flowchart for a CAT algorithm.
CATs are used for all types of formative and summative testing, however, a very
common use for CATs is to benchmark students’ progress (Challis, 2005; Irons, 2007).

Figure 2. Flowchart of CAT algorithm (Thompson & Weiss, 2011, p. 2).
The state of Utah determined that CATs are the best indicators of student progress since the assessment is able to hone in on the student’s abilities by adapting the questions to the responses that are given (Cohen, 2013). The state is in the process of creating interim benchmark CATs measured against the Utah Core Standards (Cohen, 2013). The state calls the interim benchmark the Computer-Adaptive-Assessment System (CAAS). CAAS will be offered to schools throughout the state during a window of time twice a year, fall and midyear, starting the school year 2014-2015 (USOE, 2013a).

In the course of their study, Tonidandel and colleagues (2002) discovered that CATs increase security of the assessment questions. Questions cannot be shared because each assessment is personalized to the abilities of individual students. In order to achieve a similar level of security on a conventional assessment, a parallel assessment would need to be created for each student. This study also revealed that CATs reduce assessment time. The reduced time is due to CAT’s ability to adapt to the knowledge and skill level of the student, therefore, avoiding redundant and uninformative questions. Another benefit of CAT is the instantaneous feedback (Tonidandel et al., 2002; Wainer, 2000). Instantaneous scoring and reduced assessment time can help to increase the students’ motivation.

A drawback of CAT is the lack of ability to navigate through questions. On a conventional test, students can skip questions and come back to them later or review previous questions at any time. However, in a CAT the questions must be answered in the order presented and the student may not revisit any questions (Tonidandel et al., 2002). CATs were the type of benchmark assessment used to inform the instructional
improvement cycle in the conceptual framework (Figure 1) for this study.

**Scholastic Math Inventory.** The Scholastic Math Inventory (SMI) is a CAT benchmark assessment. SMI monitors student mathematics growth and learning progress (Scholastic, 2013). The adoption of the Common Core State Standards—Mathematics (CCSS-M) brought a shift in what students are expected to learn and when. CCSS-M cites Algebra as the gateway to college and career readiness (National Governors Association [NGA] & Council of Chief State School Officers [CCSSO], 2010). According to Scholastic, SMI delivers data that directly support the practice of providing quality instruction that meets the students’ learning needs. Since all students have a right to quality and demanding mathematics curriculum (National Council of Teachers of Mathematics [NCTM], 2000), the goal of SMI is to provide data to inform the instruction that will be best for each student to succeed in mathematics, college, and career (Scholastic, 2013).

In order for SMI to provide the required data to determine the students’ learning needs, the assessment was designed to use vertical scaling by incorporating questions from below, above, and on grade level (Scholastic, 2013). Burg (2007) conducted a study to analyze the dimensional structure of mathematical achievement tests aligned to NCTM content strands (i.e., number and operation; measurement, geometry, algebra/pattern, and functions, and data analysis and probability) using several different methods for assessing dimensionality. The researcher reported three conclusions: (a) the results support the use and development of vertical scaling; (b) a lack of relationship between dimensionality and intended content strand suggests there is a connection across the five content strands,
just as NCTM intended; and (c) the connectedness of mathematic topics aides in the
determination of a potential need for early intervention by bringing in related skills to
help the students better understand the current topic.

SMI is a relatively new product. The validity research for SMI started in 2008, the
SMI program was available for purchase in 2010 (S. J. Emrich, personal communication,
February 20, 2014). The researcher was unable to locate statistical data from studies
involving the use of SMI due to the newness of the product.

**Summative Assessments**

Summative assessment is an “overview of previous learning” (Black, 1998, p. 28). Summative assessments can gather evidence over time or at the end of chapters or
phases in education (Brookhart, 2001). In Utah, the state summative test used to evaluate
students’ knowledge will be called Student Assessment for Growth and Excellence
(SAGE) starting in 2014 (USOE, 2013a). The SAGE is a high-stakes test that is state
mandated due to the pressures of NCLB Act (2002). High-stakes testing in mathematics
can have an influential power on the practices of education. Results can affect curriculum
decisions, teaching practices, school decisions, and individual students’ futures in
mathematics (Lester, 2007). Since the SAGE can determine many long lasting decisions,
it is important that the students do well on them, while still maintaining a high level of
knowledge retention. Correlation research studies suggest that when assessment literacy
is implemented through the use of proper formative assessments and feedback, the scores
on summative assessments, such as the SAGE, improve (Carlson et al., 2011; Karpinski,
2010; Keller-Margulis et al., 2008; Nugent, 2009).
Nugent (2009) conducted a study to determine a correlation between a formative benchmark assessment and a criterion-referenced summative assessment in middle school mathematics. The results indicated a strong correlation between the formative assessments and the summative assessment. Keller-Margulis and colleagues (2008) found the same strong correlation from 1,477 elementary students’ mathematics benchmark assessments and statewide end of year assessments. Karpinski (2010) examined the effectiveness of a technology-based formative assessment to predict achievement on a summative state proficiency test. The data showed that students who used technology-based formative assessment to reflect on questions and thoughtfully address the misconceptions had a positive correlation to the state test score growth. Although the school districts in these studies collected data from benchmark assessments, the districts did not have any specific process in place to use these data to make instructional decisions.

The conceptual framework diagram, displayed earlier in Chapter I (Figure 1), shows that after the formative/benchmark assessments have been obtained and the interpreted data have gone through the instructional improvement cycle, scores on the summative assessment could be predicted.

When other than correlational studies were conducted, not all research was able to find a predictive ability for summative assessments in mathematics. For example, in a quasiexperimental investigation conducted by Clariana (2009), one group of students were given wireless laptops with a 1:1 ratio of students to laptops. Another group of students had a 5:1 ratio of students to laptops. Comprehensive mathematics software was
structured around the state standards and replaced the textbooks. All students in both groups had access to the software, but the students in the group assigned laptops were allowed to go at their own pace and record keeping of students’ progress was increased. Benchmark assessments were given throughout the year. The individual laptop group outscored the group without laptops on every benchmark assessment, but not on the summative state assessment.

Similarly, Donhost (2009), in a study involving students at 86 schools, found no significant difference when he examined the predictive ability of a CAT benchmark assessment to growth on the state summative assessment for mathematics and language arts. The researcher conducted an analysis of covariance (ANCOVA) comparing the means of the summative assessments of school that used a CAT benchmark versus schools that did not use the benchmark assessments. The covariate was the summative assessment scores from a previous year. The researcher also ran a series of t-tests and a linear correlation test to determine whether the reported implementation of data-driven decision making practices (assessment literacy skills) correlated with the summative assessment scores. In both cases the null hypothesis was not rejected with an adjusted R squared = .85 for the mathematics portion of the ANCOVA.

Henderson, Petrosino, Gucken burg, and Hamilton (2007) did not find evidence to indicate a difference in mathematics scores of eighth-grade students when benchmark assessments were implemented. The researchers conducted a quasiexperimental design study using a comparative time series analysis to assess whether there was a statistically significant difference between the control and treatment schools in changes in the
mathematics summative assessment performance. They compared the scores from the students attending 22 schools using benchmark assessments with the scores from the students attending 44 schools not using benchmarks. The following year another study was performed with the benchmark assessments again being implemented the second year. Still no significant difference was found (Henderson, Petrosino, Guckenburg, & Hamilton, 2008).

Support for the use of formative assessments to improve summative assessment scores is substantial (Sherman, 2008; Wiliam et al., 2004). However, several studies do not provide sufficient evidence to support the claims made that formative benchmark assessment will improve students’ summative assessment achievement.

**Mathematical Motivation**

According to Keller (2010), instructors can purposefully design instruction and manage the learning environment to kindle student motivation. Keller developed the ARCS model of motivational design. The ARCS model is broken into four components related to motivational learning: attention, relevance, confidence, and satisfaction. Recently he expanded his model to include volition; the model now bears the acronym ARCS-V.

Keller (2010) defined each component. *Attention* is gaining, building, and sustaining the curiosity of the learner in an activity. *Relevance* is the instructor satisfying the personal needs of the learner to bring about positive results. *Confidence* is the learner’s belief that they have the ability to learn and to expect that they will be
successful based on personal efforts. **Satisfaction** is the student’s personal sense of accomplishment whether through intrinsic or extrinsic rewards. **Volition** is the self-regulated actions students take to achieve a goal. This study assumes the incorporation of each of these components as benchmark assessments are interpreted and as instruction and study habits are adjusted to achieve the goal of higher summative assessment scores and increased mathematical motivation (Figure 1).

There are many studies on student motivation, which has been a focus of educator’s efforts to improve student achievement. Teachers use an assortment of pedagogical methods to engage and motivate their students in mathematics. Motivating and engaging students can prove to be a difficult task as many intrinsic and extrinsic elements and factors can effect student mathematical motivation (Mueller et al., 2011). In general, mathematical motivation is lower than motivation in other subjects, yet, it is valued more (Plenty & Heubeck, 2013).

Weinberg and colleagues (2011) used Fishbein and Ajzen's (1975) Expectancy—Value Theory of Motivation to conduct a mixed methods study that investigated the effects of four experimental mathematics and science learning programs on the interest and motivation of middle school students. Three hundred thirty-six middle school students from seven school districts participated in a summer enrichment program that partnered with the math and science partnership from five higher educational institutions. They found that students’ interest in mathematics increased, which led to higher expectations of success. Raising the expectations of success led to higher self-efficacy in mathematics. Higher self-efficacy affected effort, persistence, and learning leading to
higher motivation to learn (Schunk, 1991, 1995, 2011). This result supports the conceptual framework for this study (Figure 1).

Mathematical motivation has been found to change for students during their years in school. In the course of a 2-year study, Plenty and Heubeck (2013) found the mathematics motivation in students from grades seven, eight, and nine changed over time. The students’ sense of mathematical importance and motivation decreased as they progressed through the school system. The study uncovered ways that teachers and students could increase mathematical motivation. One of the recommended pedagogical techniques mentioned was to focus on student assessment literacy. Plenty and Heubeck suggested that “there is a need to prevent students’ loss of faith in the usefulness of maths and to promote self-regulatory skills” (p. 28).

Peetsma and Van der Veen (2013) conducted a study with 1,168 seventh- and eighth-grade mathematics students. The researchers used questionnaires on goal orientations, self-efficacy, investment in mathematics, and well-being at school to determine the effects of performance-avoidance orientation on students’ academic motivation and achievement. The participants were low achieving students in mathematics; and at most, had performance-avoidance orientations. The researchers discovered that self-regulated learning, or assessment literacy, impacted the students’ mathematical achievement level, self-efficacy, motivation, and well-being in school. Students showing less favorable trajectories in performance-avoidance orientation also showed less favorable developments. The authors stated that “recognizing groups of students with different levels and developments of performance-avoidance orientation
makes it possible to try to intervene early in their school careers” (Peetsma & Van der Veen, 2013, p. 828).

In a combination of a cross-cultural and longitudinal qualitative study conducted by Mueller and colleagues (2011), students found mathematical motivation through positive attitudes and cooperation with peers. The researchers observed groups of students working collaboratively as they engaged in mathematical problem solving. The students helped each other justify their approach to solving problems. The peers, teachers, and the students themselves exemplified volition through their positive attitudes and a willingness to adjust problem solving approaches to obtain the desired results. The researchers identified student behaviors that showed confidence in mathematics and attention in mathematics. The students showed satisfaction in mathematics, they enjoyed doing mathematics and developed ownership of their ideas. The students also found relevance in mathematics as they discussed and share their understandings with their classmates. The study exhibited all the components of Keller’s (2010) ARCS-V model of motivational design.

Summary

In order to make assessment worthwhile, teachers and students need to become assessment literate (Popham, 2008, 2011; Sadler, 2010). A significant component of benchmark assessment data is that teachers and students understand how to use these data. This encompasses the teachers’ ability to use data from the formative assessments to adjust instruction as needed and the students’ ability to use the feedback from the
formative assessments to better their understanding of the material (Popham, 2008).

Assessments are divided into formative and summative assessments. Formative assessments comprise a range of techniques including interviews, observations, homework, and computer-based rapid assessments. The technique being examined in this study is CATs. The CAT is a form of benchmark assessments that is aligned to state standards. The SMI is a specific CAT benchmark used for this study which has the ability to deliver data that directly support the practice of providing quality instruction that meets the students’ learning needs. Summative assessments are the assessments given at the culmination of learning for a chapter, term, or year. Summative assessments are also aligned to the state standards. In Utah they will be called the SAGE.

The conceptual framework (Figure 1) is supported by research in areas of motivation, assessment, and achievement. In the field of mathematics, formative and summative assessments are given to the students. Data are collected and interpreted by both teachers and students. This study was designed to test the students’ use of data from formative assessments, in particular the SMI, to adjust instruction and study methods. The adjustments should be implemented and the instructional improvement cycle repeated. Once data have informed the instruction of the students, it could possibly predict the achievement on the summative assessment and lead to higher student motivation.

While a great deal of research has been done on formative assessment (Black & Wiliam, 1998a; Black et al., 2003), relatively few studies have involved middle-school-aged students. Research findings led to developments of strategies designed to maximize
student potentials. Benchmark assessments are one of these strategies. Proper use of data provided by benchmark assessments can help teachers adjust the instruction and students adjust study habits, ultimately resulting in increased summative assessment scores (Carlson et al., 2011; Karpinski, 2010; Keller-Margulis et al., 2008; Nugent, 2009). Formative benchmark assessments inform teachers and students of gaps in knowledge. Informed adjustments to teachers’ instruction and students’ study habits increase students’ self-efficacy. Elevated self-efficacy increases intellectual performance and construction of knowledge, which, in theory, makes the students’ mathematical motivation soar.
CHAPTER III
METHODOLOGY

The purpose of this study was to determine the impact of formative SMI computer adaptive benchmark assessments and assessment literacy training on seventh- and eighth-grade student mathematics achievement and motivation. Student achievement and student motivation were the dependent variables. The independent variables were the assessment literacy training and SMI benchmark assessments.

This chapter describes the research methodology used in the quantitative study. The chapter presents the research questions, design, participants, setting, instrumentation, procedures, data analysis, and interpretation for the study.

Research Questions

1. Do formative SMI computer adaptive benchmark assessments, coupled with assessment literacy training for students, influence seventh- and eighth-grade students’ performance on summative assessments in mathematics?
   a. Null hypothesis: The use of SMI and assessment literacy training does not make a difference in student performance on summative assessments in mathematics.

2. Does training in benchmark assessment literacy increase student motivation?
   a. Null hypothesis: There is no difference in the means scores for motivation between the seventh- and eighth-grade students who received training in benchmark assessment literacy and those in the control groups, as
measured by the Instructional Materials Motivation Survey.

Design

Quantitative research in educational settings can be used to indicate valid instructional methods and curricula (Schoenfeld, 2000). This quantitative study used a posttest, quasiexperimental design. Creswell (2008) defined a quasiexperimental study as an experimental design where the treatment and control groups do not have random participant assignment; often they are groups of convenience. The quasiexperimental groups used in this study were groups of convenience. The groups were preexisting classes. This design was used to determine if use of SMI and assessment literacy makes a difference in mathematical achievement and student motivation for seventh- and eighth-grade students. The students who did not receive the treatment had the unit taught to them in the same manner as the treatment groups, except they were not trained in assessment literacy and did not receive the SMI benchmark assessments. A posttest and a motivation survey were administered to measure the difference between the two groups after treatment (Heffner, n.d.).

Participants

Forty-four seventh-grade students and 34 eighth-grade students from one charter school participated in this study, representing all the seventh- and eighth-grade students except for the two honors classes. The researcher, who was also the teacher, had access to the students at the charter school. Therefore, a convenience grouping was used to create a
control group and an experiment group. Each group contained 22 seventh-grade and 17 eighth-grade students. All students were selected based upon (a) enrollment in one of the participating mathematics classes, (b) student willingness to participate in the study, and (c) receipt of the student’s signed parental consent form.

Although the researcher was unable to randomly assign all students for the study, efforts were made by the administration to ensure that each class was representative of the school population. When student class schedules were created, student placement in mathematics was considered before scheduling any other classes. This made the task of placing an equivalent number of high, average, and low performing students in each class and having the number of males and females equally distributed.

The total group of participants involved 33 males and 45 females. About 92% of the participants were Caucasian, the other 8% were of varying ethnicities. Seventeen percent of the student body was on free or reduced lunch. Approximately 10% of the students had individual education plans (IEP).

Setting

The study took place in an IB charter school that served families from suburban communities in Northern Utah. The school served 675 students from kindergarten to eighth grade, of which 140 were seventh- and eighth-grade students. The charter school was a Title 1 school with 17% eligible for free or reduced lunch. Ninety-two percent of the total student body was Caucasian, 4% were Asian/Pacific Islander, 3% were Hispanic, and 1% were African American. Out of this population of students, less than
1% received English as a second language services and 10% met the eligibility requirements for special education services. The student body was comprised of 49.4% males and 50.5% females.

The mathematics curriculum used by the charter school was McGraw-Hill’s *My Math* for kindergarten through fifth grade and McGraw-Hill/Glencoe’s *Glencoe Math: Your Common Core Edition* for sixth grade through eighth grade. Teachers were required to teach the Utah State Mathematics Core as well as incorporate the IB standards into the curriculum.

During the study, the SMI assessments were administered to the students in the mathematics classroom. The 25 desks were moved around to form five tables. Twenty-five laptop computers were set up on the tables with five computers on each table. The link to the SMI assessment was written on the middle of the whiteboard at the front of the classroom so all the students could read it to log into the SMI assessment. The students’ individual passwords were printed on adhesive name tags and distributed. Every student had their own computer and they worked independently. The posttest and all instructional activities were also conducted in the mathematics classroom. The mathematics classroom was a carpeted room with a teacher’s desk in one corner and 25 student desks and chairs arranged in groups of two or three, all facing the front wall where the whiteboard was located. On the whiteboard, a Mimio® bar was attached to transform the whiteboard into an interactive board where most of the instruction took place. One instructor, the researcher, taught all the participants and administered the assessments, including the SMI assessments.
Material

The classroom lessons and the posttest used *Glencoe Math: Your Common Core Edition CCSS* textbook (Carter et al., 2013) as a resource. The lessons were mostly lecture and practice with inquiry and hands-on labs interspersed throughout the unit. The labs used material such as graph paper, graphing calculators, IPads, and other manipulatives to assist the students in discovering connections among rate of change, slopes, intercepts, and linear equations.

Data Sources/Instruments

Four instruments were utilized during the data collection stage of this study. The instruments were all designed to measure different aspects of the students’ academic performance and the students’ motivation in mathematics. They included the Utah Criterion-Referenced Test (CRT), SMI Benchmark Assessment, the Equations in Two Variables unit posttest, and the Instructional Materials Motivation Survey (IMMS).

Criterion-Referenced Tests

The CRT is a standardized test comprised of three core curriculum sections; English Language Arts, Mathematics, and Science. According to the Joint Committee on Testing Practices (2004), all standardized tests must meet psychometric (test study, design, and administration) standards for reliability, validity, and lack of bias. The CRT is a standardized test and considered to be reliable and valid (USOE, 2013b). The researcher used the CRT scores from 2013 as a covariate since the CRT scores were
readily available to the researcher.

**Scholastic Math Inventory Benchmark Assessment**

The benchmark assessment that was used for the treatment group to test for content knowledge readiness was designed and developed by SMI. SMI is a computer-adaptive test that presents items to the students according to the answer given on the previous item. The teacher provided the SMI assessment program with the student’s grade level. SMI then began by presenting the student with an item that has a skill measured at the proficient level for the student’s grade. SMI increased the difficulty level after the student provided a set of correct answers and decreased the difficulty level when wrong answers were provided. This process was continued until the skill level of the student had been narrowed. The assessment concluded after the student had answered 25 to 30 questions or had gone past 40 minutes since the assessment began. In this study the SMI benchmark assessment tested to see if the students had adequate previous knowledge to continue learning new concepts found in the Equation in Two Variables unit taught to the seventh grade and eighth grade classes.

SMI was created and validated in 2004 by MetaMetrics, Inc., an educational measurement and research organization. SMI was developed to use The Quantile Framework® for Mathematics. The Quantile Framework is a scientifically based system that links assessment to instruction, in this case, the Utah State Mathematics Core. Scores from SMI provided the teacher and the students with a “quantile score” that showed the level of understanding measured against the Utah State Mathematics Core (Scholastic,
Using the standards found in the Utah State Mathematics Core, SMI had predetermined the proficient level for each grade, seventh-grade ranged from 890 to 1010 while the eighth-grade ranged from 1020 to 1140. The data provided a breakdown of the different standards in which the students needed interventions to be prepared to learn the next concepts.

**Equations in Two Variables Unit Posttest**

The Equations in Two Variables unit posttest (Carter et al., 2013) was used to measure achievement (see Appendix A). This curriculum-based assessment focused on the content that the researcher presented during the unit. The questions for the Equations in Two Variables unit assessment were designed by *Glencoe Math: Your Common Core Edition CCSS* (Carter et al., 2013). The students had 40 minutes to complete a 15-question multiple-choice test.

**Instructional Materials Motivation Survey**

The Instructional Materials Motivation Survey (IMMS; Keller, 2010) was used to measure student motivation (see Appendix B). Permission was obtained from the survey’s author to modify and use in this study (see Appendices C and D). The survey was administered to both the control and treatment groups at the end of the study to determine if there was a difference in the students’ motivation between the control and treatment groups.

The IMMS used a Likert-type scale to measure responses of the participants. The
response scale ranged from 1 to 5, with 1 signifying “not true” and a 5 signifying “very true.” There were 36 items on the IMMS that measured the motivational effects of the instructional treatment. Each item fell into one of four categories: attention (12 items), confidence (nine items), relevance (nine items), and satisfaction (six items). The survey allowed for individual scoring of each category or an overall score (Keller, 2010).

Reliability testing of the IMMS survey, using Cronbach’s alpha measure, indicated that each of the four categories and the overall scores had satisfactory reliability coefficients which were between .81 and .96 (Keller, 2010). Keller stated the validity of IMMS was established preparing two sets of instructional material. Both lessons had the same objectives and content. The lesson for the control group was prepared to the standards of the course, but was not enhanced to make it interesting. The lesson for the experimental group “was enhanced with strategies to stimulate curiosity, illustrate the practical relevance of the content, build confidence, and provide satisfying outcomes” (Keller, 2010, p. 286). The participants were randomly assigned and the lessons were completed in one class period. The scores for the experimental group were much higher than the control group. According to Keller, the IMMS can be adapted to fit various situations for differing research needs.

**Procedures**

This study used a quasiexperimental design to determine if there was a difference in the mean scores between treatment and control groups on a unit posttest. The treatment groups received the SMI benchmark assessments and training in assessment literacy. The
control groups were taught the unit in the same manner as the treatment groups, but without the SMI benchmark assessments and training in assessment literacy. At the conclusion of the unit, both the treatment group and the control group were measured for mathematical achievement using a unit posttest and data were analyzed using ANCOVA to determine if there was a difference in the mathematical achievement between the groups. The treatment and control groups were also measured for mathematical motivation using the IMMS. Analysis of IMMS scores using Analysis of Variance (ANOVA) determined if there was a difference in mathematical motivation between the treatment groups and the control groups.

The researcher, who was also the instructor, conducted the research in a period of seven weeks. The remainder of the procedure section will explain the process of obtaining permissions to conduct this study. It will then describe the training that was received by the teacher and the students. The section will conclude with an explanation of how the two research questions were answered.

Permissions

An application to the Institutional Review Board (IRB) of Utah State University was submitted for the research study and permission was granted (see Appendix E). Permission was obtained from the charter school’s Head of School (see Appendix F) and the School Board to perform the study (see Appendix G). Permission for participation in the study was requested and obtained from parents of the student (see Appendix H). No one was considered a participant and no data were collected without prior written consent from the participant and the parent or legal guardian. All participants were made aware
that they could opt out of the research at any time.

Students and parents were reassured that all data gathered during this study would not reveal any student’s identity. They were also reassured that if the results suggested that there does indeed exist a benefit to the treatment, all students would be given the same treatment after the study was completed.

Training

The researcher/teacher had a phone conversation with the Scholastic trainer to discuss the needs for training. The Scholastic trainer then provided educational material to the teacher through email communication. The training material described how to administer SMI and how to interpret the data. The material also included suggestions on how to use textbooks to address gaps in student mathematical skills. However, the *Glencoe Math: Your Common Core Edition* textbook (Carter et al., 2013) had not been published long enough to be included in the lists of textbook resources provided by Scholastic. The teacher read through the training material and attempted to practice administering the SMI to lower grade students. However, Scholastic’s technology and the charter school’s technology did not sync together. After one week of the technology specialists working together, it was decided that the charter school would load the SMI program onto their school’s computer server. The researcher/teacher was unable to administer a practice assessment. During the first administration of the SMI to the treatment group it was discovered that when the system was transferred to the school’s server, some of the students’ login information had been deactivated. The logins were reactivated within five minutes and the SMI administration was successful (see Appendix
I). The SMI training material can be accessed through the following websites:
http://teacher.scholastic.com/math-assessment/scholastic-math-inventory,
http://www.quantiles.com, and http://teacher.scholastic.com/math-assessment/scholastic-
math-inventory/ekit_thankyou.htm.

Assessment literacy training for the students took place 2 weeks prior to the start
of the Equations in Two Variables unit. The students in the treatment group received
training in becoming assessment literate. They received instruction on how to read and
interpret data results from the SMI benchmark assessments as well as unit assignments
and quizzes. The teacher used the information gained from scholastic’s training material
to teach the students how to use the data to adjust their study habits. Examples of
assessment scores were provided for the students to practice assessment literacy skills.
Students were observed with sample assessment scores to ascertain their assessment
literacy. The students were considered assessment literate if they were able to interpret
the scores and find a resource to close the gaps in knowledge. This process took place in
the classroom during class time and lasted three days, including administration of the first
SMI benchmark assessment (see Appendix J). During this time, students in the control
group were being utilized by second and third grade teachers to assist their students in
reading and mathematics.

Question # 1

Question #1: Do formative SMI computer adaptive benchmark assessments,
coupled with assessment literacy training for students, influence seventh- and eighth-
grade students’ performance on summative assessments in mathematics?
Initial SMI benchmark assessment. The SMI benchmark assessments must be administered no less than one month apart. In order for the second SMI assessment to be administered during the middle of the unit, the first SMI assessment was given two weeks before the unit began and directly after the assessment literacy training ended. The SMI benchmark assessment is a computer adaptive test that adjusts the questions presented to each student according to their previous answers. The purpose of the SMI was to determine if the students have enough background knowledge to learn the new concepts in the mathematics unit on Equations in Two Variables. The scores were recorded and both the teacher and the students received the results from the SMI. The teacher and the students took the data from the SMI to assist in determining which concepts needed extra emphasis for the students to learn the new information presented in the unit. The SMI data would list the concept areas that were lower than expected for a student in their grade level. The students and the teacher would discuss what sections the teacher’s data inferred needed to be retaught. Then the students would use their individual data to find concept areas that were different than the overall class needs. The students’ assignment after receiving the data was to find a section in their math book or online that would address their personal need. The students were held accountable for this task by turning in to the teacher some evidence of completing the assignment. Most of the students turned in completed review assignments.

The teachers and administration who worked in the charter school in which the study took place were aware of the assessment schedule for the treatment group. The second and third grade teachers requested to have the control group use this time to assist
the younger students read and do mathematics.

**Instruction time and second SMI benchmark assessment.** Two weeks after the administration of the first SMI benchmark assessment, instruction for the Equations in Two Variables unit took place. The unit took 5 weeks to teach and was separated into two sections. Data from the first SMI assessment were used by both the teacher and students to adjust instruction and students’ approach to study. None of the treatment group students received data indicating they already knew the material that was to be taught in this unit. The teacher recognized from the data that the students in the treatment group needed extra emphasis on rational expressions. As the new concepts were being taught, the teacher would show more examples of linear equations using fractions and decimals. Occasionally, the teacher would spend some student work time to reteach how to get common denominators or how to convert a fraction to a decimal or vice versa.

At the end of a two and a half week period another SMI benchmark assessment was taken by the students. The SMI was administered in the classroom where each student had access to their own laptop computer to take the assessment. The scores were recorded and both the teacher and the students received the results from the SMI. The teacher and the students took the data from the SMI to assist in determining which concepts needed extra emphasis. A few of the students found they still needed to work on the same mathematical concepts as the first SMI data showed. Some of the students saw improved SMI scores and were able to recognize a new set of mathematical concepts. Other students saw a decrease in scores and discovered they were farther behind in mathematical content knowledge. The teacher saw that a portion of the class still needed
work on rational expression, but decided to spend the next two and a half weeks putting an emphasis on modeling and computing with integers. During this time, the students in the control groups were administered a mid-chapter quiz from the textbook, which was a normally scheduled activity for students when not receiving the SMI treatment.

**Instruction time and summative assessment.** The data received from the second SMI assessment were used again by both the teacher and the students to make needed adjustments. The remaining concepts of the unit were taught over the next two and a half weeks. The teacher made a point to include extra problems and examples that had the students modeling and computing with integers. At the conclusion of the unit, a summative assessment was administered to both the treatment and control groups. The purpose of the summative assessment was to measure content knowledge the students learned during the unit. Data from the summative assessment in the Equations in Two Variables unit was analyzed using ANCOVA with the initial group differences in mathematics achievement being the covariant.

The dependent variables were the results of the unit posttest. The independent variables were the assessment literacy training combined with the SMI benchmark information that the students used to adjust their study habits for the treatment group and no training combined with the SMI data for the control group. The covariate variable was the 2013 CRT mathematics proficiency scores for all the participants. The covariate was used to equalize the pre-existing mathematical abilities of the students. Prior to the ANCOVA test the participants were tested for linearity, homogeneity of regression, normality, homogeneity of variance, and reliability of the covariate.
Question #2

Question #2: Is there a difference in motivation mean scores between seventh and eighth grade mathematics students who received training in assessment literacy and knowledge of the data received from the SMI benchmark assessments as measured by the Instructional Materials Motivation Survey (IMMS)?

At the conclusion of the Equations in Two Variables unit and after the summative assessment was given, the IMMS was administered to all participants in the study. This was done with the aid of Google Drive Forms online during class time in the computer lab (see Appendix K). The data gathered did not contain any individual names or other identifying information. Using the online survey caused a nonthreatening environment where the students could answer the questions truthfully without the worry of losing their anonymity.

Electronic data collection is becoming popular with quantitative researchers (Creswell, 2008). It provides a quick and easy way to collect data. Researchers believe that using the confidential format of online surveys may help prevent apprehension and encourage honest replies (Ary, Jacobs, Sorensen, & Walker, 2013; Whelan, 2008). IMMS helped determine the mathematical motivation using an application of the general linear model (GLM), called ANOVA.

Data Analysis

This study investigated the impact SMI benchmarks and assessment literacy had on summative assessments and student motivation in mathematics. The researcher
analyzed unit posttest scores for seventh- and eighth-grade mathematics control and treatment groups. The researcher also analyzed data received from the IMMS. The parametric tests used for these analyses are discussed in this section.

The General Linear Model

Both hypotheses were tested using applications of the GLM. The GLM incorporates a number of different statistical models including the tests used with this study, ANCOVA and ANOVA. The general linear model is a generalization of multiple linear regression models when more than one dependent variable is used. It is a flexible model in which the errors are usually assumed to follow a multivariate normal distribution, normally distributed and independent. If the errors do not follow a multivariate normal distribution, generalized linear models may be used to relax assumptions using various techniques (Salkind, 2011).

Analysis of Student Achievement Data

An ANCOVA was conducted on data gathered from the Equations in Two Variables unit posttest to see if there was a difference in scores of students who received SMI benchmarks and assessment literacy training and students who did not receive this treatment. Participants’ CRT scores from 2013 were utilized as a covariate to statistically adjust initial group differences in mathematics achievement as part of the ANCOVA test. Salkind (2011) stated that ANCOVA “allows you to equalize initial differences between groups” (p. 299). The analysis covaried on math achievement overall.

According to Warner (2012), in order for the results from the ANCOVA to be
utilized and considered generalizable for the intended population, several assumptions must be met: linearity, homogeneity of regression, normality, homogeneity of variance, and reliability of the covariate. Prior to the ANCOVA test, these assumptions were tested and met.

**Analysis of Student Motivation Data**

Data were gathered from the IMMS that measured mathematical motivation in students. The IMMS was comprised of four categories (attention, relevance, confidence, and satisfaction) that were correlated, as determined by previous research (Keller, 2010). ANOVA was an appropriate statistical method for this study since it is a flexible statistical model that relates responses to linear predictor variables (Warner, 2012). ANOVA tested if there was a difference between the IMMS scores for the students who received the SMI benchmarks and assessment literacy training versus the students who did not receive this treatment.

In order for results to be utilized and generalized to the population of interest, several assumptions were verified using appropriate statistical analyses in IBM SPSS version 20. These assumptions were: the independence, normality and homogeneity of the variances.

**Interpretation**

IBM SPSS Version 20 was used to run the statistical analyses for this study. A Sensitivity power analysis, using G*Power 3.1, alpha = .05, N = 78. The sample size was large enough to detect differences between group means at the 95% confidence level.
Summary

This quantitative study was conducted in a public IB charter school with seventh- and eighth-grade students during a seven week period of time in mathematics. During the course of the study, two treatment groups (one seventh-grade class and one eighth-grade class) received SMI benchmark assessments and training in assessment literacy, while the two control groups (one seventh-grade class and one eighth-grade class) were taught the unit in the usual manner. At the conclusion of the unit, both the treatment and control groups were measured for mathematical achievement and motivation using *Glencoe Math: Your Common Core Edition* posttest and the IMMS. The posttest was used to determine the difference in mathematical achievement. Student posttest scores were collected and analyzed using the ANCOVA. The IMMS scores were used to determine if there was a difference in motivation between the treatment groups and the control groups. The scores from the IMMS were collected and analyzed using ANOVA. Table 1 describes the timeline for the treatment group procedures of this study. Table 2 describes the timeline for the control group procedures.

Table 1
*Treatment Group Procedure Timeline*

<table>
<thead>
<tr>
<th>Activity</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assessment Literacy Training/Initial SMI Benchmark Assessment</td>
<td>3 days</td>
</tr>
<tr>
<td>Time between training and Equations in Two Variables unit spent completing previous unit</td>
<td>2 weeks</td>
</tr>
<tr>
<td>Equations in Two Variables unit instruction begins</td>
<td>2 weeks</td>
</tr>
<tr>
<td>Second SMI Benchmark Assessment</td>
<td>1 day</td>
</tr>
<tr>
<td>Equations in Two Variables unit instruction continues</td>
<td>2 weeks</td>
</tr>
<tr>
<td>Unit Posttest administered</td>
<td>1 day</td>
</tr>
<tr>
<td>IMMS administered</td>
<td>1 day</td>
</tr>
</tbody>
</table>
Table 2

*Control Group Procedure Timeline*

<table>
<thead>
<tr>
<th>Activity</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assist second- and third-grade students read and do mathematics</td>
<td>3 days</td>
</tr>
<tr>
<td>Time between training and Equations in Two Variables unit spent completing previous unit</td>
<td>2 weeks</td>
</tr>
<tr>
<td>Equations in Two Variables unit instruction begins</td>
<td>2 weeks</td>
</tr>
<tr>
<td>Mid-Chapter review</td>
<td>1 day</td>
</tr>
<tr>
<td>Equations in Two Variables unit instruction continues</td>
<td>2 weeks</td>
</tr>
<tr>
<td>Unit Posttest administered</td>
<td>1 day</td>
</tr>
<tr>
<td>IMMS administered</td>
<td>1 day</td>
</tr>
</tbody>
</table>
CHAPTER IV

RESULTS

The purpose of this study was to determine the impact of SMI computer adaptive benchmark assessments and assessment literacy training on seventh- and eighth-grade students’ mathematics achievement and motivation. The study also investigated student motivation caused by assessment literacy and knowledge of the data received from the SMI benchmark assessments.

Chapter IV is organized into five sections: (a) demographic information for the participants; (b) research questions and hypotheses; (c) data analysis and results of an ANCOVA that measured the impact of SMI and assessment literacy training on students’ mathematics achievement compared to that of students without the treatment; (d) data analysis and results of ANOVAs that measured the impact of SMI and assessment literacy training on students’ mathematics motivation compared to that of students without the treatment; and (e) a summary of the results.

Demographics

Forty-four seventh-grade students and 34 eighth-grade students from one charter school participated in this study, representing all the seventh- and eighth-grade students except for the two honors classes. A convenience grouping was used to create a control group and an experiment group. Each group contained 22 seventh-grade and 17 eighth-grade students. These students were selected based upon (a) enrollment in one of the participating mathematics classes, (b) student willingness to participate in the study, and
(c) receipt of the student’s signed parental consent form. All students completed both the unit posttest and the IMMS survey. Table 3 presents the demographic information for the participants in terms of ethnicity. Table 4 presents the gender of the participants by instructional groups. Table 5 presents the grade demographics by instructional groups.

**Research Questions and Hypotheses**

For this study, the researcher examined the impact of formative SMI computer adaptive benchmark assessments and assessment literacy training on seventh- and eighth-grade student mathematics achievement and motivation in one charter middle school.

**Table 3**

*Ethnicity of Students by Instructional Group*

<table>
<thead>
<tr>
<th>Ethnicity</th>
<th>Treatment group</th>
<th>Control group</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
</tr>
<tr>
<td>African American</td>
<td>1</td>
<td>2.56</td>
<td>0</td>
</tr>
<tr>
<td>Asian</td>
<td>2</td>
<td>5.13</td>
<td>0</td>
</tr>
<tr>
<td>Caucasian</td>
<td>36</td>
<td>92.31</td>
<td>36</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0</td>
<td>0.00</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>39</td>
<td>100.00</td>
<td>39</td>
</tr>
</tbody>
</table>

**Table 4**

*Gender of Students by Instructional Group*

<table>
<thead>
<tr>
<th>Gender</th>
<th>Treatment group</th>
<th>Control group</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
</tr>
<tr>
<td>Male</td>
<td>15</td>
<td>38.46</td>
<td>18</td>
</tr>
<tr>
<td>Female</td>
<td>24</td>
<td>61.54</td>
<td>21</td>
</tr>
<tr>
<td>Total</td>
<td>39</td>
<td>100.00</td>
<td>39</td>
</tr>
</tbody>
</table>
Table 5

*Grade of Students by Instructional Group*

<table>
<thead>
<tr>
<th>Grade</th>
<th>Treatment group</th>
<th>Control group</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
</tr>
<tr>
<td>Seventh</td>
<td>22</td>
<td>56.41</td>
<td>22</td>
</tr>
<tr>
<td>Eighth</td>
<td>17</td>
<td>43.59</td>
<td>17</td>
</tr>
<tr>
<td>Total</td>
<td>39</td>
<td>100.00</td>
<td>39</td>
</tr>
</tbody>
</table>

There are two questions that guided this research.

1. Do formative SMI computer adaptive benchmark assessments, coupled with assessment literacy training for students, influence seventh- and eighth-grade students’ performance on summative assessments in mathematics?

   a. Null hypothesis: The use of SMI and assessment literacy training does not make a difference in student performance on summative assessments in mathematics.

2. Does training in benchmark assessment literacy increase student motivation?

   a. Null hypothesis: There is no difference in the means scores for motivation between the seventh- and eighth-grade students who received training in benchmark assessment literacy and those in the control groups, as measured by the IMMS.

   **Data Analysis and Results: Question 1—Student Achievement**

   **ANCOVA**

   An ANCOVA was conducted using IBM SPSS Version 20 to determine the
impact of formative Scholastic Math Inventory (SMI) computer adaptive benchmark assessments and assessment literacy training on seventh- and eighth-grade student mathematics achievement. Student achievement was the dependent variable. The independent variables were the assessment literacy training and SMI benchmark assessments. The 2013 CRT Mathematics scores for each student served as the covariate.

**ANCOVA Descriptive Statistics**

The 2013 CRT mathematics scores were obtained for the 78 students participating in this study and were used as the covariate. The sample size, means, and standard deviations of the covariate for the two instructional groups are listed in Table 6. These descriptive statistics show very little difference in the means and standard deviations of the covariate in the two groups.

A posttest was given to both instructional groups at the end of the unit to test student mathematics achievement, the dependent variable after the treatment group received SMI assessment benchmarks and assessment literacy training. The sample size, means, and standard deviations of the dependent variable for the two instructional groups are listed in Table 7. Descriptive statistics show the means for the control group are

Table 6

*Descriptive Statistics for Math CRT Percent Scores, the Covariate, by Instructional Group*

<table>
<thead>
<tr>
<th>Instructional group</th>
<th>n</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control group</td>
<td>39</td>
<td>85.92</td>
<td>3.85</td>
</tr>
<tr>
<td>Treatment group</td>
<td>39</td>
<td>85.62</td>
<td>3.49</td>
</tr>
<tr>
<td>Total</td>
<td>78</td>
<td>85.77</td>
<td>3.65</td>
</tr>
</tbody>
</table>
Table 7

Descriptive Statistics for Unit Posttest Percent Scores, the Dependent Variable, by Instructional Group

<table>
<thead>
<tr>
<th>Instructional group</th>
<th>n</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control group</td>
<td>39</td>
<td>47.01</td>
<td>16.54</td>
</tr>
<tr>
<td>Treatment group</td>
<td>39</td>
<td>44.79</td>
<td>15.06</td>
</tr>
<tr>
<td>Total</td>
<td>78</td>
<td>45.90</td>
<td>15.75</td>
</tr>
</tbody>
</table>

higher and standard deviations indicate more variation for the control group than the treatment group

**ANCOVA Assumptions**

Preliminary analyses were conducted to evaluate assumptions for the ANCOVA. Those assumptions include: (a) reliability of covariate, (b) linearity, (c) homogeneity of regression, (d) normality, and (e) homogeneity of variance.

Reliability of the covariate, CRT mathematics score, was assumed due to the information provided by the Joint Committee on Testing Practices (2004). According to the Joint Committee on Testing Practices, all standardized tests must meet psychometric (test study, design, and administration) standards for reliability, validity, and lack of bias.

One of the assumptions for ANCOVA is a linear relationship between the covariate (CRT scores) and the dependent variable (unit posttest scores). The linear relationship is shown in Figure 3 by examining the scatterplot of unit posttest scores versus CRT math scores. The Pearson coefficient, \( r \), of 0.54, \( p < .05 \) indicates a moderate positive linear relationship between unit posttest scores and CRT math scores. This accounts for 29% of variance of success on the unit posttest by success on the CRT.
The regression slopes must be equal for each group (Brace, Kemp, & Snelgar, 2012; Salkind, 2011). The assumption of homogeneity of regression was tested by including the interaction term in the general linear model. The interaction of SMI assessments with assessment literacy training and CRT math scores yielded $F(1,78) = 1.05, p = .31$, and $\eta^2 = 0.01$. These data indicated that the interaction of SMI assessments with assessment literacy training and CRT math scores was not statistically significant at $\alpha = 0.05$, and therefore did not violate the homogeneity of regression assumption for ANCOVA.

**Figure 3.** Posttest percent scores versus CRT mathematics percent scores.
The researcher used the Shapiro-Wilk test to determine if the assumption of normality was met. Table 8 provides the results for the Shapiro-Wilk normality test for the posttest scores of the instructional groups. This table indicates there is no statistically significant evidence at the $\alpha = .05$ significance level, meaning the assumption of normality had been met.

The final assumption for ANCOVA tested by the researcher was homogeneity of variance, meaning that the distribution of the dependent variable for the two groups being compared must have constant variance across factor levels. The results showed no statistical significance of the homogeneity of variances, $F(1,76) = .98, p = .33$; therefore, equal variances can be assumed.

**ANCOVA Results**

A one-way between-subjects analysis of covariance was carried out to assess the impact the SMI assessment benchmarks and assessment literacy training had on the overall performance on the unit posttest for seventh- and eighth-grade mathematics students. Checks were carried out to confirm homogeneity of regression and linear relationship between the covariate and dependent variables. The between-subjects factor

<table>
<thead>
<tr>
<th>Instructional group</th>
<th>Statistic</th>
<th>df</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>.95</td>
<td>39</td>
<td>0.09*</td>
</tr>
<tr>
<td>Treatment</td>
<td>.95</td>
<td>39</td>
<td>0.07*</td>
</tr>
</tbody>
</table>

*Normality assumed, $p > .05$. 

Table 8

*Shapiro-Wilk Test of Normality for the Posttest of the Control and Treatment Groups*
comprised of two groups: the treatment group who had received the SMI assessment benchmarks and assessment literacy training and a control group of mathematics students. The covariate comprised the 2013 CRT Mathematics scores, and this was related to the overall unit posttest scores; \( F(1, 75) = 30.42, p < .0005, \) partial \( \eta^2 = .29 \).

Adjusting for this covariate the ANCOVA for control versus treatment on the unit posttest scores was found to have no statistically significant effects: \( F(1, 75) = .26, p > .05, \) partial \( \eta^2 < .01 \). Table 9 shows the ANCOVA results for the unit posttest scores with the CRT mathematics scores as the covariate.

These results indicate that instruction, which utilized SMI and assessment literacy training, did not produce a statistically significant difference in posttest scores. The observed power of .80 indicates that if the parameters are as expected, the researcher would reject the null hypothesis 80% of the time when this experiment is conducted.

The adjusted mean score for the treatment group was 45.13 compared to 46.67 for the control group. Table 10 displays the adjusted means for the posttest scores when using the CRT Mathematics score of 85.77 to evaluate the model. From this perspective, it is possible to see the difference in means for the control group and the treatment group.

Table 9

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
<th>( \eta^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRT covariate</td>
<td>1</td>
<td>5485.98</td>
<td>30.42</td>
<td>0.00</td>
<td>0.29</td>
</tr>
<tr>
<td>Group posttest</td>
<td>1</td>
<td>45.99</td>
<td>0.26</td>
<td>0.62</td>
<td>0.00</td>
</tr>
<tr>
<td>Error</td>
<td>75</td>
<td>180.36</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\( R^2 = .292 \) (Adjusted \( R^2 = .273 \))
Table 10

*Adjusted Means and 95% Confidence Interval for Unit Posttest Scores*

<table>
<thead>
<tr>
<th>Instructional group</th>
<th>$M$</th>
<th>$SE$</th>
<th>Lower bound</th>
<th>Upper bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>46.67</td>
<td>2.15</td>
<td>42.38</td>
<td>50.95</td>
</tr>
<tr>
<td>Treatment</td>
<td>45.13</td>
<td>2.15</td>
<td>40.84</td>
<td>49.42</td>
</tr>
</tbody>
</table>

There was a 1.54 point difference in the adjusted means, which showed a smaller disparity than the 2.22 point difference in the actual mean posttest scores between classes. Note that the control group had a higher mean and adjusted means than that of the treatment group.

After interpreting the results of the ANCOVA analysis, the null hypothesis was not rejected based upon the result of no statistically significant difference in means.

**Data Analysis and Results: Question 2—Student Motivation**

**ANOVA**

An ANOVA was conducted for each of the four subscales and the overall score of the IMMS (attention, relevance, confidence, satisfaction, and overall score) to determine the effect SMI assessments and assessment literacy training had on student motivation. The ANOVA was used to assess the differences between two or more group means when there is one independent variable and one dependent variable. The independent variable was the SMI assessment and assessment literacy training. The dependent variable was the student motivation.
ANOVA Descriptive Statistics

The IMMS survey was administered to the 78 students participating in this study. The sample size, means, and standard deviations of each of the subscales for the two instructional groups are listed in Table 11. The statistics show the means and standard deviations of the treatment group were higher than the control group in all subscales and the overall score.

ANOVA Assumptions

Preliminary analyses were conducted to evaluate assumptions for the ANOVA. Those assumptions include: (a) the independence, (b) normality, and (c) homogeneity of variance.

The researcher attempted to ensure that the assumption of independence was met by receiving training on how to proctor the IMMS survey for the treatment and control groups. Classroom control and a quiet environment were maintained during administration of the online survey in all classes. The researcher circulated among the

Table 11

Descriptive Statistics for IMMS Subscales and Overall Score by Instructional Group

<table>
<thead>
<tr>
<th>Subscale</th>
<th>Control group</th>
<th>Treatment group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
</tr>
<tr>
<td>Attention</td>
<td>3.07</td>
<td>0.67</td>
</tr>
<tr>
<td>Relevance</td>
<td>3.38</td>
<td>0.60</td>
</tr>
<tr>
<td>Confidence</td>
<td>3.30</td>
<td>0.70</td>
</tr>
<tr>
<td>Satisfaction</td>
<td>3.18</td>
<td>0.81</td>
</tr>
<tr>
<td>Overall score</td>
<td>3.22</td>
<td>0.57</td>
</tr>
</tbody>
</table>
students as they completed the survey and provided uniform directions. Within the class and across classes, students did not talk or discuss the survey questions. Thus, the researcher assumed that each participant’s scores on the IMMS variables were independent from scores of all the other participants.

The researcher used the Shapiro-Wilk test to determine if the assumption of normality was met for each of the IMMS subscales. Table 12 provides the results for the Shapiro-Wilk normality test for the instructional groups in each of the IMMS subscales. This table indicates there is no statistically significant evidence at the $\alpha = .05$ significance level for both the control and treatment groups in the subscales of attention, relevance, confidence, overall score, and the treatment group in satisfaction, meaning the assumption of normality was met for all but the control group in the subscale of satisfaction. The appearance of a moderate departure from normality does not necessarily imply a serious violation of the assumptions. Montgomery (2009) stated, “Because the $F$

Table 12

<table>
<thead>
<tr>
<th>Subscale</th>
<th>Control group</th>
<th>Treatment group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Statistic</td>
<td>df</td>
</tr>
<tr>
<td>Attention</td>
<td>0.95</td>
<td>39</td>
</tr>
<tr>
<td>Relevance</td>
<td>0.96</td>
<td>39</td>
</tr>
<tr>
<td>Confidence</td>
<td>0.98</td>
<td>39</td>
</tr>
<tr>
<td>Satisfaction</td>
<td>0.94</td>
<td>39</td>
</tr>
<tr>
<td>Overall score</td>
<td>0.97</td>
<td>39</td>
</tr>
</tbody>
</table>

*Normality assumed, $p > .05$. 
test is only slightly affected, we say that the analysis of variance is robust to the normality assumption” (p. 77).

The data must also meet the assumption of homogeneity of variance, meaning that the distribution of the dependent variable for the groups being compared must have constant variance across factor levels. Levene’s test of equality of variances was completed using IBM SPSS. Table 13 shows the results of the test. Based on the results of the Levene’s Test of Equality of Variances, the researcher found no statistically significant evidence that the variances across the groups were unequal. Therefore, homogeneity of variance can be assumed.

**ANOVA Results**

An ANOVA was conducted for each of the four subscales and the overall score of the IMMS (attention, relevance, confidence, satisfaction, and overall score) to compare the effects SMI assessments and assessment literacy training had on student motivation. Table 14 shows the ANOVA results for the effects of SMI benchmark assessments and assessment literacy training on the five dependent variables. There were no significant

<table>
<thead>
<tr>
<th>Table 13</th>
<th>Levene’s Test of Equality of Variances</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subscale</td>
<td>$F$</td>
</tr>
<tr>
<td>Attention</td>
<td>2.30</td>
</tr>
<tr>
<td>Relevance</td>
<td>1.82</td>
</tr>
<tr>
<td>Confidence</td>
<td>.01</td>
</tr>
<tr>
<td>Satisfaction</td>
<td>.27</td>
</tr>
<tr>
<td>Overall score</td>
<td>1.28</td>
</tr>
</tbody>
</table>

*Equal variance assumed $p > .05$. 
Table 14

One-Way Analysis of Variance for the Effects of SMI Benchmark Assessments and Assessment Literacy Training on Four Subscales and the Overall Score

<table>
<thead>
<tr>
<th>Variable and source</th>
<th>SS</th>
<th>MS</th>
<th>$F(1,76)$</th>
<th>$p$</th>
<th>$\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attention</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between</td>
<td>0.45</td>
<td>0.45</td>
<td>0.90</td>
<td>.35</td>
<td>.01</td>
</tr>
<tr>
<td>Within</td>
<td>37.85</td>
<td>0.50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relevance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between</td>
<td>0.46</td>
<td>0.46</td>
<td>1.08</td>
<td>.30</td>
<td>.01</td>
</tr>
<tr>
<td>Within</td>
<td>32.77</td>
<td>0.43</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Confidence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between</td>
<td>1.31</td>
<td>1.31</td>
<td>2.63</td>
<td>.11</td>
<td>.03</td>
</tr>
<tr>
<td>Within</td>
<td>37.85</td>
<td>0.50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Satisfaction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between</td>
<td>3.21</td>
<td>3.21</td>
<td>4.88</td>
<td>.03*</td>
<td>.06</td>
</tr>
<tr>
<td>Within</td>
<td>50.07</td>
<td>0.66</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between</td>
<td>0.96</td>
<td>0.96</td>
<td>2.39</td>
<td>.13</td>
<td>.03</td>
</tr>
<tr>
<td>Within</td>
<td>30.39</td>
<td>0.40</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$p < .05$

effects at the $p < .05$ level of the SMI benchmark assessment and assessment literacy on student motivation in attention, $F(1, 76) = 0.90, p = .35$, partial $\eta^2 = .01$; relevance, $F(1, 76) = 1.07, p = .30$, partial $\eta^2 = .01$; confidence, $F(1, 76) = 2.63, p = .11$, partial $\eta^2 = .03$; and overall score, $F(1, 76) = 2.39, p = .13$, partial $\eta^2 = .03$. However, there was a significant effect at the $p < .05$ level of the SMI benchmark assessment and assessment literacy on student motivation in satisfaction, $F(1, 76) = 4.88, p = .03$, partial $\eta^2 = .06$. Based on the results, there was insufficient evidence to reject the null hypothesis. However, there was sufficient evidence found to show an increase in the satisfaction subscale of motivation, as measured by the IMMS.
Summary

Through the use of this quasiexperimental, posttest research study, the researcher determined there was no statistical difference between student achievement scores and the level of student motivation for seventh- and eighth-grade students who received formative SMI computer adaptive benchmark assessments and assessment literacy training.

Statistical analyses were performed to determine the effects of the formative SMI computer adaptive benchmark assessments and assessment literacy training. An ANCOVA was performed using the unit posttest scores as the dependent variable and the 2013 CRT mathematics scores as the covariate to determine the effect the treatment had on student achievement. An ANOVA was conducted on each of the subscales and the overall score to determine the effect the treatment had on student motivation.

The use of formative SMI computer adaptive benchmark assessments and assessment literacy training resulted in no statistically significant difference in student achievement. There was not sufficient evidence to reject the null hypothesis for question 1: The use of SMI and assessment literacy training does not make a difference in student performance on summative assessments in mathematics.

The student motivation scores for the satisfaction subscale indicated a statistically significant difference between the control and treatment groups. However, results from the ANOVA for the other three subscales (attention, relevance, and confidence) and the overall scores indicated there was not sufficient evidence to reject the null hypothesis for question 2: There is no difference in the means scores for motivation between the
seventh- and eighth-grade students who received training in benchmark assessment literacy and those in the control groups, as measured by the IMMS. Chapter V will discuss the implications of the results of the use of formative SMI computer adaptive benchmark assessments and assessment literacy training on student achievement and motivation.
CHAPTER V
DISCUSSION

Chapter IV presented data analyses utilizing ANCOVA and ANOVA to compare the differences between mathematical achievement and motivation scores of seventh- and eighth-grade students who received formative SMI computer adaptive benchmark assessments and assessment literacy training versus those who did not receive the treatment in one charter middle school. Descriptive statistics and summaries were also presented.

The purpose of this chapter is to review and discuss the findings of this study. This chapter is divided into the following sections: a summary of the study, which will restate the problem and purpose; a discussion of the findings and implication; revisit the theoretical framework; the limitations of the study; recommendations for further research; and the conclusion.

Summary of the Study

NCLB (2002) mandated that all students make academic growth and improve student achievement. The focus of attention from states, districts, and schools has been on the methods and strategies used to measure student achievement (Popham, 2004). Assessments for accountability, or summative assessments, are designed to summarize the achievement level of students. In order to assist students in showing academic growth on their summative assessments, formative assessments are used to inform the instructional decision making process (Popham, 2010).
The problem addressed in this study was that many teachers and students are lacking the ability to use data from formative mathematical assessments to adjust their instruction and study habits to improve student summative assessment results in mathematics (Havnes, 2004; Popham, 2008; Stiggins, 2007). This ability is known as assessment literacy. A review of the literature affirmed conflicting results about the predictive ability formative mathematics assessments have on summative mathematics assessment scores (Carlson et al., 2011; Clariana, 2009; Donhost, 2009; Henderson et al., 2008; Karpinski, 2010; Keller-Margulis et al., 2008; Nugent, 2009). Another problem addressed in this study was that self-efficacy and motivation is diminished among middle school students who perform poorly on high stakes tests (Bandura, 1977).

The purpose of this study was to determine the impact of SMI computer adaptive benchmark assessments and assessment literacy training on seventh- and eighth-grade students’ mathematics achievement. The study also investigated student mathematical motivation caused by assessment literacy and knowledge of the data received from the SMI benchmark assessments.

**Discussion**

**Research Question 1: Student Achievement**

The finding of this study supported the findings of other researchers who found no significant difference in scores when comparing students who were administered CAT benchmark assessments and students who were not administered the CAT assessments (Clariana, 2009; Donhost, 2009; Henderson et al., 2007, 2008). For two consecutive
years, Henderson and colleagues (2007, 2008) were not able to find evidence to indicate a difference in mathematics scores of eighth-grade students when benchmark assessments were implemented in a study similar to this study.

The findings of this study were contradictory to several correlational research studies that suggested when assessment literacy is implemented through proper use of formative benchmark assessments and feedback, the scores on summative assessments improved (Carlson et al., 2011; Karpinski, 2010; Keller-Margulis et al., 2008; Nugent, 2009). The findings of this study not only contradicted several other studies by not finding a significant difference in the unit posttest, but the control group posttest scores were higher than the treatment group posttest, even after adjusting for the covariate. There are many possible explanations for the lack of significant evidence to reject the null hypothesis, or the fact the control group posttest scores were higher. Explanations could include: that the unit was too short to give sufficient time for the study habit adjustments to begin to make a difference; SMI did not provide the proper information to inform the teacher’s instruction or the students’ study habits; the assessment literacy training was insufficient; or a Type II error. These explanations may suggest further research.

A review of the literature supported the use of formative benchmark assessments and assessment literacy training for both the teacher and the students. The review also found study results that were not conclusive about whether the use of formative benchmark assessments actually improved student mathematical achievement. The knowledge gained through the review of literature and the current study’s findings
suggested that administrators and educators who want to improve student mathematical achievement may want to investigate further into the use SMI formative benchmark assessments. Bull and McKenna (2003) pointed out that the cost of the assessment program and the time it takes to implement the technology are disadvantages and limitations of assessing with technology. The results of this study also suggested that administrators may want to reconsider whether use of computer adaptive benchmark assessments are worth the expense.

The current study contributed to the field of existing research by adding a quantitative study on the impact of SMI computer adaptive benchmark assessments and assessment literacy training on seventh- and eighth-grade students’ mathematics achievement. Most available research studies on the topic of SMI computer adaptive benchmark assessments were related to the validation of the criterion referenced assessment. No research was found on the impacts of using SMI. Most available research on topics related to formative assessment and summative assessment did not include the use of assessment literacy training.

**Research Question 2: Student Motivation**

The findings of this study indicated that the students in the treatment group who received assessment literacy training and were administered the SMI benchmark assessments were positively impacted in the satisfaction component of student mathematical motivation. However, the attention, relevance, and confidence components and the overall score were not found to be statistically significant. This contradicts results from the author of the IMMS (Keller, 2010), who found the four components to be
correlated. Although the attention, relevance, confidence, and the overall scores were not considered statistically significant, mean scores for those subscales and overall scores for the treatment group were higher than those for the control group. This could imply that a longer study was needed to fully identify the differences between the means.

According to Keller (1987, 2010), it is possible for teachers to identify motivational needs and to design enhancements that will improve student performance and motivation. Using Keller’s ARCS model can help teachers purposefully design instruction to increase student motivation (Keller, 2010). The ARCS model of motivational instruction was designed to assist teachers in creating an environment that acquires and maintains attention, help students understand the personal relevance of the material, help build and maintain student confidence, and to ascertain satisfaction (Keller, 2010).

The researcher did not specifically use Keller’s model of instructional design in this study. However, the teacher was already using many of Keller’s (1987, 2010) suggestions for improving motivation in the four different areas. The teacher would ask the students questions that would generate the learner’s curiosity; for example, the teacher asked the students what types of careers they were interested in possibly doing someday. For each career, the teacher and the students would attempt to find a benefit for the knowledge of Equations in Two Variables. Every career had a use for this knowledge, like recognizing that over time their pay should increase and the students should be able to predict how long it might take them to earn the salary they wanted. Keller’s (1987, 2010) suggestions for improving attention included piquing the learner’s curiosity by
generating inquiry. This tactic used by the teacher was similar to another suggestion made by Keller (1987, 2010) for increasing relevance, which was to give concrete examples to the students’ current interests. To build motivation through confidence in a student, Keller (1987, 2010) indicated that providing frequent feedback can build confidence. The teacher took extra care to incorporate frequent and detailed feedback for the students in the treatment group. Keller (1987, 2010) suggested that allowing learners to share or affirm their efforts can build motivation through satisfaction. The teacher would occasionally have the students work in groups on problem sets, then allow someone from each group to share how their group solved a certain problem.

Keller’s (2010) model of instructional design was not addressed in this study’s questions, therefore, was not the focus. However, the IMMS posed questions that aided in answering the study’s question of increased motivation. Some examples of the survey items that aided in answering the question were: “There was something interesting at the beginning of this unit that got my attention”; “There were stories, pictures, or examples that showed me how the material in this unit could be important to some people”; and “The content of this material is relevant to my interests.” The researcher believes that if the study had been designed to make use of the full ARCS model for instruction, the results for the attention, relevance, and confidence components of the IMMS may have been different.

In general, motivation is lower in mathematics than in other subjects (Plenty & Heubeck, 2013). Researchers posited the decline in mathematical motivation may be effectively addressed by enhancing students’ assessment literacy and study skills (Black
& Wiliam, 1998a; Hattie et al., 1996; Heritage, 2007; McDonald & Boud, 2003; Nicol, 2009; Plenty & Heubeck, 2013; Popham, 2008; Sadler, 1998; Smith et al., 2013; Stiggins, 2007). Research has shown that teaching students assessment literacy and study skills, and providing feedback, does not guarantee the student will gain mathematical motivation (Plenty & Heubeck, 2013; Zimmerman & Martinez-Pons, 1992).

Theoretical Framework Revisited

This project was based on Paul Black and Dylan Wiliam’s theoretical framework of Assessment for Learning. Black and Wiliam (2009) posited that the responsibility for learning lies equally with the teacher and the student. The teacher is responsible for designing and implementing an effective learning atmosphere. The student is responsible for the learning that goes on within that environment (Black & Wiliam, 2009; Leahy et al., 2005; Sadler, 1998). By understanding where responsibilities lie for learning, Black and Wiliam were able to develop five key strategies that conceptualize the theory of Assessment for Learning:

1. clarifying and sharing learning intentions and criteria for success;
2. engineering effective classroom discussions and other learning tasks that elicit evidence of student understanding;
3. providing feedback that moves learners forward;
4. activating students as instructional resources for one another; and
5. activating students as the owners of their own learning. (pp. 4-5)

In this study, the researcher/teacher was diligent in presenting objectives and rubrics to provide opportunities for student success. Lessons were designed to allow students to work in groups and have classroom discussions, along with assignments to gain evidence of student understating. The SMI was administered and feedback was
given to the students in order to recognize areas in their math content knowledge that needed to be improved. The students were expected to use the feedback to assist in finding the proper resources that would help their learning move forward.

The results suggested that the theory on the impact of benchmarks and assessment literacy on student achievement and motivation, as found in Figure 1, should be reconsidered. Possibly the SMI, which was meant to inform the teacher and the students of what adjustments needed to be made in instruction and study habits, did not provide data useful to make the required adjustments.

The theoretical framework of *Assessment for Learning* provided opportunities to challenge learners to reflect on their own thinking. It required students to be self-regulated learners by being assessment literate and using the feedback from formative assessment to adjust their learning (Black & Wiliam, 2009; Stiggins & Chappuis, 2006). *Assessment for Learning* implied that through benchmark feedback and assessment literacy, students’ mathematical achievement and, subsequent, motivational increase could be obtained or predicted. The results of this study found no evidence to support this theory.

**Limitations**

This study needed to be cautious of generalization due to limitations and constraints. The research was limited to a small sample size of seventh- and eighth-grade mathematics students in a public charter school with the lack of diversity in the school’s student demographics (e.g., the student population is 92% Caucasian). Students in a
different grade level, at a different school, in a different geographical location, or in a different subject area could have different academic strengths and weaknesses.

The participants in this study were not randomly assigned to a group, as classes were already intact before the research began. While the researcher was not allowed to randomly assign the students for the study, school administration randomly placed students into classes prior to the beginning of the school year. Each class was assigned an equivalent number of high, average, and low performing students. Students with Individual Learning Plans through the resource department were placed in classes according to their individual needs. Efforts were made by administration to ensure similar demographics of each group.

The study had the constraint that the research was done on only one mathematical unit instead of the whole course. When a whole course is studied, different results could occur due to different strengths in content knowledge in each section taught throughout the course. The unit chosen for this study could have been more difficult than other units, which could affect the mathematical achievement and mathematical motivation of the students.

The 2013 CRT mathematics score was used as a covariate to control for previous math achievement; there were several factors that were not taken into account. Factors such as additional math support, level of parental support and involvement, extracurricular activities, family responsibilities, gender, ethnicity, or socioeconomic status were not included in the statistical analysis. None of these factors were relevant to the research questions.
The use of surveys has inherent limitations. Online surveys must be completed in one sitting. The students may have rushed through the survey if they felt they would not have enough time to complete it in class. Surveys rely on self-reported answers to questions. Inaccuracies in data collection can occur with poor memory, misunderstanding of the question, or answering with what the student thinks the researcher wants to see. The students were encouraged to answer each question as honestly as possible. The IMMS used a Likert-type scale with five possible answers, creating the possibility of a midpoint system threat (Swain, Weathers, & Neidrich, 2008).

The teacher, who was also the researcher, taught students in both the treatment and control groups. Several suggestions were made to alleviate validity threats to research (Downing, 2004). To enhance the validity of this study, a journal was kept by the researcher explaining how instructor bias was dealt with while teaching the unit. Data from the first SMI helped the teacher identify that the treatment group needed extra instruction working with rational expressions. The teacher “found it very difficult to give extra examples and problems with fractions and decimals to the treatment group and not to the control group, even though they are sure to need the extra help as well.” Yet, the teacher created two separate lesson plans for the two groups and did not present any extra help to the control group. When the students would come in for extra help after school with their homework, the teacher took extra precautions to not give the control group students the same help with the rational numbers as the treatment group students. Before the teacher began the after school assistance, the student’s name was verified on the list as to which group they belonged to so the teacher would not inadvertently give a member
of the control group the type of examples to aide in their assignment that was reserved for the treatment group. The same precautions were made after identifying the need for extra work on integers after the second SMI assessment. The teacher recalls, “There was one time I was helping [Student A], from the treatment group, and [Student B], from the control group, came in the room. [Student B] actively participated in solving the problem I started helping [Student A] with.” This situation was unavoidable, yet, added to the limitations and constraints for this study.

**Recommendation**

Analyses of the data showed no statistical evidence that the SMI formative benchmark assessment coupled with assessment literacy training made a difference in student mathematical achievement and little difference in student mathematical motivation. However, a review of the literature suggests there may still be benefits to using formative benchmark assessments (Black & Wiliam, 2009b; Goertz et al., 2009; Harlen & Crick, 2003; Karpinski, 2010; Keller-Margulis et al., 2008; Nugent, 2009; Riggan & Oláh, 2011; Shapiro & Gebhardt, 2012; Stiggins & Chappuis, 2006; Tonidandel et al., 2002) and to train teachers and students in assessment literacy (Blink, 2007; Brown et al., 2009; Christman et al., 2009; Gibbs & Simpson, 2004; Hattie et al., 1996; Havnes, 2004; Heritage, 2007; Hoover, 2009; Ingram et al., 2004; McDonald & Boud, 2003; Nicol, 2009; Oláh et al., 2010; Popham, 2008, 2011; Sadler, 1998; Shapiro & Gebhardt, 2012; Shepard et al., 2011; Smith et al., 2013; Stiggins, 1995; Wayman & Stringfield, 2006). Based on the results from this study, the researcher recommends that
Recommendations for Further Research

Results from this study showed that the mathematical achievement means of the control group was higher than the treatment group. The study also indicated that only one of the four components of the ARCS model for motivational learning showed improvement from the project intervention. A mixed method research would have helped to discover the reasons for these unexpected statistics. The current study should be replicated with the addition of observations to discover the interaction students have with assessment data and study habits and interviews with both groups of students could explore the students’ understanding on the content knowledge, the SMI data, and mathematical motivation. Assessment data could be analyzed for mean differences between control groups and treatment groups. Qualitative data rich in information gathered from individual students could be combined with quantitative statistical data to discover how and if benchmark assessments and assessment literacy improves student mathematical achievement and mathematical motivation.

Recommendations for further research would also include research into the CAAS benchmark assessment program offered by the USOE. The state of Utah is in the process of creating an interim benchmark CAT measured against the Utah Core Standards (Cohen, 2013). The state calls the interim benchmark the Computer-Adaptive-Assessment System (CAAS). CAAS will be offered to schools throughout the state during a given window of time twice a year, fall and midyear, starting the school year
2014-2015 (USOE, 2013a). Research into CAAS could discover if it provides usable data to inform teachers’ instruction to improve student mathematical achievement. Research could be conducted to determine the effects of student assessment literacy and study skills on students’ mathematical motivation. Further research could be conducted to isolate the variables that might have impact on student mathematical achievement and mathematical motivation. These variables could be found inside the classroom, such as, the type of instruction given (e.g., inquiry based learning) or the feeling of acceptance the student has from the teacher. These variables could also be found outside the classroom, such as, the level of parental support, family responsibilities, or education level of the parents.

**Conclusion**

The results of this study indicated that seventh- and eighth-grade students who received instruction for the Equations in Two Variables unit along with SMI benchmark assessments combined with assessment literacy training showed no improvement in mathematics achievement. The findings from this study demonstrated that student mean scores for mathematical motivation improved on one of the four components of the ARCS model for motivational learning. The results showed that the student mathematical motivation overall mean scores were not statistically significant. Thus, this study was unable to reject the null hypothesis: The use of SMI and assessment literacy training does not make a difference in student performance on summative assessments in mathematics, and there is no difference in the means scores for motivation between the seventh- and
eighth-grade students who received training in benchmark assessment literacy and those in the control groups, as measured by the IMMS.
REFERENCES


Burg, S. S. (2007). *An investigation of dimensionality across grade levels and effects on vertical linking for elementary grade mathematics achievement tests.* Retrieved from http://books.google.com/books?hl=en&lr=&id=wkRpl8RBT5UC&oi=fnd&pg=PR12&dq=%22consequences+of+violating+the+assumption+of+unidimensionality+h+e+22%22+to+which+evidence+and+theory+support+the+interpretations+of+test+scores,+it+is+22%22+sources+of+multidimensionality+may+exist+particularly+in+a%22+&ots=V7NlFSK3M-&sig=T6F7OtCVgXmrPFVkEwouHnckadk


APPENDICES
Appendix A

Equations in Two Variables Unit Test
Equations in Two Variables Test

Write the letter for the correct answer in the blank at the right of each question.

1. What is the constant rate of change between \( x \) and \( y \) in the table below?

<table>
<thead>
<tr>
<th>( x )</th>
<th>1</th>
<th>5</th>
<th>9</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td>( y )</td>
<td>-6</td>
<td>-3</td>
<td>0</td>
<td>3</td>
</tr>
</tbody>
</table>

A. \(-\frac{4}{3}\)  
B. \(-\frac{3}{4}\)  
C. \(\frac{3}{4}\)  
D. \(\frac{4}{3}\)  

1. __________

2. What is the slope of the line that passes through the points \( A(-2, -1) \) and \( D(3, 5) \)?

F. \(\frac{6}{5}\)  
G. \(\frac{5}{6}\)  
H. \(-\frac{5}{6}\)  
I. \(-\frac{6}{5}\)  

2. __________

3. What are three numbers that have a sum of 35 if the greatest number is 14 more than the least number?

A. 6, 7, 21  
B. 5, 11, 19  
C. 10, 11, 24  
D. 1, 15, 15  

3. __________

4. The costs of cookies at store A are shown in the graph. The cost \( y \) for \( x \) cookies at store B is represented by the equation \( y = 0.30x \).
Which of the following statements is true?

F. The cookies at store A cost more.  
G. The cookies at store A cost \$0.50 each.  
H. The cookies at store B cost \$0.15 each  
I. The cookies at store B cost more.  

4. __________

5. What are the slope and \( y \)-intercept for the graph of \( y - 7x = 10 \)?

A. slope: 7, \( y \)-intercept: 10  
B. slope: 7, \( y \)-intercept: -10  
C. slope: -7, \( y \)-intercept: 10  
D. slope: -7, \( y \)-intercept: -10  

5. __________

6. Which is the equation in slope-intercept form for the graph of the line shown?

F. \(y = -3x - 2\)  
G. \(y = -3x + 2\)  
H. \(y = 3x - 2\)  
I. \(y = 3x + 2\)  

6. __________
7. David is having his birthday party at a water park. The park charges $150 plus $10 per guest. The total cost of the party $y$ can be represented by the equation $y = 10x + 150$. What does the slope represent?

A. the number of guests  
B. the cost to rent the water park  
C. the cost per guest  
D. David's age

8. Which equation, in point-slope form, passes through (3, -1) and has a slope of 2?

F. $y + 1 = 2(x - 3)$  
G. $y - 1 = 2(x + 3)$  
H. $y + 1 = 2(x + 3)$  
I. $y - 1 = 2(x - 3)$

9. What are the x- and y-intercepts for the graph of $2x - 5y = 10$?

A. x-intercept: -5, y-intercept: 2  
B. x-intercept: -5, y-intercept: -2  
C. x-intercept: 5, y-intercept: -2  
D. x-intercept: 5, y-intercept: 2

10. Xavier has $20 more than Sara. Their combined money totals $90. Which system of equations represents this situation?

F. $x + s = 90$  
G. $x + s = 90$  
H. $x - s = 90$  
I. $s - x = 90$

11. Which of the following is the solution of the system of equations shown?

A. (2, 2)  
B. (-2, 2)  
C. (2, -2)  
D. (-2, -2)

12. What is the solution of the system of equations?

\[ \begin{align*}
  y &= x - 4 \\
  y &= -3x \\
F. (3, -1) & G. (-3, 1) & H. (-1, 3) & (1, -3)
\end{align*} \]

13. What is the solution of the system of equations?

\[ \begin{align*}
  y &= x - 10 \\
  y &= 2x + 5 \\
\end{align*} \]
14. What is the constant rate of change shown in the graph below?

A. 2.5      B. 3      C. 7.5      D. 10

15. What is the slope of the line from the table below?

<table>
<thead>
<tr>
<th>Hour</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance</td>
<td>50</td>
<td>100</td>
<td>200</td>
<td>400</td>
</tr>
</tbody>
</table>

F. 50      G. 25      H. \( \frac{1}{25} \)      I. \( \frac{1}{50} \)
Appendix B

Instructional Materials Motivation Survey (IMMS)
Instructional Materials Motivation Survey

There are 36 statements in this questionnaire. Please think about each statement in relation to the instructional materials you have just studied and indicate how true it is. Give the answer that truly applies to you, and not what you would like to be true, or what you think others want to hear.

Think about each statement by itself and indicate how true it is. Do not be influenced by your answers to other statements.

Record your responses on this Google Drive Form. Thank you.

Use the following values to indicate your response to each item.

1 = Not true  
2 = Slightly true  
3 = Moderately true  
4 = Mostly true  
5 = Very true

1. When I first looked at this unit, I had the impression that it would be easy for me.
2. There was something interesting at the beginning of this unit that got my attention.
3. This unit was more difficult to understand than I would like for it to be.
4. After reading the introductory information, I felt confident that I knew what I was supposed to learn from this unit.
5. Completing the exercises in this unit gave me a satisfying feeling of accomplishment.
6. It is clear to me how the content of this unit is related to things I already know.
7. Many of the pages in this unit had so much information that it was hard to pick out and remember the important points.
8. These materials in the textbook for this unit are eye-catching.
9. There were stories, pictures, or examples that showed me how the material in this unit could be important to some people.
10. Completing this unit successfully was important to me.
11. The quality of the examples and problems helped to hold my attention.
12. This unit had very little in it that captured my attention.
13. As I worked on this unit, I was confident that I could learn the content.
14. I enjoyed this unit so much that I would like to know more about this topic.
15. The pages of this unit look dry and unappealing.
16. The content of this material is relevant to my interests.
17. The way the information is arranged on the pages of the textbook for this unit helped keep my attention.
18. There are explanations or examples of how people use the knowledge in this unit.
19. The exercises in this unit were too difficult.
20. This unit has things to encourage my curiosity.
21. I really enjoyed studying this unit.
22. The amount of repetition in this unit caused me to get bored sometimes.
23. The content and style of writing in this unit convey the impression that its content is worth knowing.
24. I learned some things that were surprising or unexpected.
25. After working on this unit for a while, I was confident that I would be able to pass a test on it.
26. This unit was not relevant to my needs because I already knew most of it.
27. The wording of feedback after the assignments and quizzes, or of other comments in this unit, helped me feel rewarded for my effort.
28. The variety of assignments and examples, helped keep my attention on the unit.
29. The style of writing and teaching is boring.
30. I could relate the content of this unit to things I have seen, done, or thought about in my own life.
31. There are so many words on each page that it is irritating.
32. It felt good to successfully complete this unit.
33. The content of this unit will be useful to me.
34. I could not really understand quite a bit of the material in this unit.
35. The good organization of the content helped me be confident that I would learn this material.
36. It was a pleasure to work on such a well-designed unit.
Appendix C

Approval Email from IMMS Author, John Keller, to Use/Modify IMMS Survey
Dear Sheryl,

Thank you for contacting me. Your study sounds interesting and I am happy that you will incorporate a modified version of the IMMS. You are welcome to do so.

Best wishes,
john keller

John M. Keller, Ph.D.
Professor Emeritus
Educational Psychology and Learning Systems
Florida State University

9705 Waters Meet Drive
Tallahassee, FL 32312-3746
Phone: 850-294-3908

Official ARCS Model Website: http://arcsmodel.com. UPDATED 18 SEP 2013
Professional Website: http://mater.fsu.edu/~jekeller/JohnsHome/
Appendix D

Modifications of Instructional Material Motivation Survey (IMMS)
Instructional Materials Motivation Survey

There are 36 statements in this questionnaire. Please think about each statement in relation to the instructional materials you have just studied and indicate how true it is. Give the answer that truly applies to you, and not what you would like to be true, or what you think others want to hear.

Think about each statement by itself and indicate how true it is. Do not be influenced by your answers to other statements.

Record your responses on the answer sheet that is provided and follow any additional instructions that may be provided in regard to the answer sheet that is being used with this survey [this Google Drive Form. Thank you.]

Use the following values to indicate your response to each item.

1 = Not true
2 = Slightly true
3 = Moderately true
4 = Mostly true
5 = Very true

1. When I first looked at this lesson [unit], I had the impression that it would be easy for me.
2. There was something interesting at the beginning of this lesson [unit] that got my attention.
3. This material [unit] was more difficult to understand than I would like for it to be.
4. After reading the introductory information, I felt confident that I knew what I was supposed to learn from this lesson [unit].
5. Completing the exercises in this [unit] gave me a satisfying feeling of accomplishment.
6. It is clear to me how the content of this material [unit] is related to things I already know.
7. Many of the pages [in this unit] had so much information that it was hard to pick out and remember the important points.
8. These materials [in the textbook for this unit] are eye-catching.
9. There were stories, pictures, or examples that showed me how the material in this unit could be important to some people.
10. Completing this lesson [unit] successfully was important to me.
11. The quality of the writing [examples and problems] helped to hold my attention.
12. This lesson [unit] is so abstract that it was hard to keep [had very little in it that captured] my attention on it.
13. As I worked on this lesson [unit], I was confident that I could learn the content.
14. I enjoyed this lesson [unit] so much that I would like to know more about this topic.
15. The pages of this lesson [unit] look dry and unappealing.
16. The content of this material is relevant to my interests.
17. The way the information is arranged on the pages [of the textbook for this unit] helped keep my attention.
18. There are explanations or examples of how people use the knowledge in this lesson [unit].
19. The exercises in this lesson [unit] were too difficult.
20. This lesson [unit] has things that stimulated [to encourage] my curiosity.
21. I really enjoyed studying this lesson [unit].
22. The amount of repetition in this lesson [unit] caused me to get bored sometimes.
23. The content and style of writing in this lesson [unit] convey the impression that its content is worth knowing.
24. I learned some things that were surprising or unexpected.
25. After working on this lesson [unit] for a while, I was confident that I would be able to pass a test on it.
26. This lesson [unit] was not relevant to my needs because I already knew most of it.
27. The wording of feedback after the assignments and quizzes, or of other comments in this lesson [unit], helped me feel rewarded for my effort.
28. The variety of reading passages, exercises, illustrations, etc. [assignments and examples], helped keep my attention on the lesson [unit].
29. The style of writing and teaching is boring.
30. I could relate the content of this lesson [unit] to things I have seen, done, or thought about in my own life.
31. There are so many words on each page that it is irritating.
32. It felt good to successfully complete this lesson [unit].
33. The content of this lesson [unit] will be useful to me.
34. I could not really understand quite a bit of the material in this lesson [unit].
35. The good organization of the content helped me be confident that I would learn this material.
36. It was a pleasure to work on such a well-designed lesson [unit].
Appendix E

Institutional Review Board Approval
Institutional Review Board
USU Assurance:
FWA#00003308
Expedite #7
Letter of Approval

FROM: Melanie Domenech Rodriguez,
IRB Chair

True M. Rubal, IRB Administrator

To: James Dorward, Sheryl Rushton

Date: September 25, 2013

Protocol #: 5360

Title: The Impact Of Computer-Adaptive Benchmark Data And Assessment Literacy On Student Achievement And Motivation In Mathematics

Risk: Minimal risk

Your proposal has been reviewed by the Institutional Review Board and is approved under expedite procedure #7 (based on the Department of Health and Human Services (DHHS) regulations for the protection of human research subjects, 45 CFR Part 46, as amended to include provisions of the Federal Policy for the Protection of Human Subjects, November 9, 1998):

Research on individual or group characteristics or behavior (including, but not limited to, research on perception, cognition, motivation, identity, language, communication, cultural beliefs or practices, and social behavior) or research employing survey, interview, oral history, focus group, program evaluation, human factors evaluation, or quality assurance methodologies. This approval applies only to the proposal currently on file for the period of one year. If your study extends beyond this approval period, you must contact this office to request an annual review of this research. Any change affecting human subjects must be approved by the Board prior to implementation. Injuries or any unanticipated problems involving risk to subjects or to others must be reported immediately to the Chair of the Institutional Review Board.

This approval applies only to the proposal currently on file for the period of one year. If
your study extends beyond this approval period, you must contact this office to request an annual review of this research. Any change affecting human subjects must be approved by the Board prior to implementation. Injuries or any unanticipated problems involving risk to subjects or to others must be reported immediately to the Chair of the Institutional Review Board.

Prior to involving human subjects, properly executed informed consent must be obtained from each subject or from an authorized representative, and documentation of informed consent must be kept on file for at least three years after the project ends. Each subject must be furnished with a copy of the informed consent document for their personal records.

4460 Old Main Hill Logan, UT 84322-4460

PH: (435) 797-1821
Fax: (435) 797-3769
WEB: irb.usu.edu
EMAIL: irb@usu.edu
Appendix F

Approval by Head of School to Conduct Research
September 30, 2013

To Whom It May Concern:

I, Heather Shepherd, grant permission to allow Sheryl Rushton to conduct a quantitative study examining the impact of computer-adaptive benchmark data and assessment literacy on student achievement and motivation in middle school mathematics to be conducted at Channing Hall. I understand that the information gathered would be for research purposes only, and the identity and identifying information of all participants will be kept confidential.

Sincerely,

Heather Shepherd
Head of School
Channing Hall
Appendix G

Approval by the School Board to Conduct Research
September 30, 2013

To Whom It May Concern:

I, Rachal Milford, grant permission to allow Sheryl Rushton to conduct a quantitative study examining the impact of computer-adaptive benchmark data and assessment literacy on student achievement and motivation in middle school mathematics to be conducted at Channing Hall. I understand that the information gathered would be for research purposes only, and the identity and identifying information of all participants will be kept confidential.

Sincerely,

Rachal Milford
Board President
Channing Hall
Appendix H

Student Participant Consent Form
Student Participant Consent Form

Introduction: Your child is being asked to participate in a research project for a doctoral dissertation for Sheryl Rushton under the direction of Dr. James Dorward in the Department of Education at Utah State University. Your child was selected as a possible participant because they fit the criteria for this study. There will be approximately 90 participants in this research.

Procedures: One 7th grade math class and one 8th grade math class will receive a benchmark assessment, Scholastic Math Inventory (SMI), given at the beginning of a unit and again half way through the unit. They will also receive training in assessment literacy teaching them how to interpret assessment results. Another 7th grade math class and another 8th grade math class will be taught without the use of SMI assessments and assessment literacy training, being taught with the teacher’s normal procedures. As your child’s teacher, I will administer an expert-validated end-of-unit assessment covering the standards taught in one unit of math instruction. The results will be compared between the experimental groups and the control groups. At the end of the unit, your child will be asked to complete an online student survey consisting of approximately 36 questions intended to determine student motivation in their math class. The participants will be participating in the research for the duration of seven weeks.

Benefits: If the findings from the research determine benefits exist from receiving the treatment then all students will receive the treatment for the remainder of the school year.

Explanation and offer to answer questions: Please ask me, Sheryl Rushton, any questions you have to help you understand this research project. I will gladly answer any inquiries regarding the purpose and procedures of the present study. Inquiries or questions regarding this research may be directed to (801) 572-2709 or my email address srushton@channinghall.org.

Voluntary nature of participation and right to withdraw without consequence: Refusal to participate in this study will have no effect on your child’s grade or any future services they may be entitled to from the school system. Should you agree to participate in this study and decide later that you wish to withdraw; you will be free to withdraw from the study at any time without penalty. If you agree to participate at this time, please sign and date this statement. Thank you very much for your willingness to participate in this research project.

Confidentiality: I will take precautions to protect participant identity by not using the names of participants or a specific class period in my results or writing. The survey will be located on a secure online server. Data is stored on the server and is kept in a password-protected database. The information will be stored on the survey site for the duration of three years and then I will delete it. I will store all research documentation on a password-protected Dropbox account for the duration of three years and will then delete it.
the documentation from the computer database. Any hard copies of the data will be stored in a locked filing cabinet and shredded at the end of three years. A video recording will be made of my lessons training the students in assessment literacy only if a student is absent for the lesson. No visual image of the students will be seen on the video.

**IRB Approval Statement:** The Institutional Review Board for the protection of human participants at Utah State University has approved this research study. If you have any questions or concerns about your rights or a research-related injury and would like to contact someone other than the researcher, you may contact the IRB Administrator at (435) 797-0567 or email irb@usu.edu to obtain information or to offer input.

**Copy of consent:** You have been given two copies of this Informed Consent. Please sign both copies and keep one copy for your files.

**Investigator Statement:** “I certify that the research study has been explained to the individual, by me, and that the individual understands the nature and purpose, the possible risks and benefits associated with taking part in this research study. Any questions that have been raised have been answered.”

**Signature of Researcher:**

__________________________________________
Sheryl Rushton  
Principal Investigator  
(801)-573-7937  
rushtonsj@comcast.net

**Child/Youth Assent:** I understand that my parent(s) or guardian(s) are aware of this research study and that they have given permission for me to participate. I understand that it is up to me to participate even if they say yes. If I do not want to be in this study, I do not have to and no one will be upset if I don’t want to participate or if I change my mind later and want to stop. I can ask any questions that I have about this study now or later. By signing below, I agree to participate.

**Student Name (printed):** __________________________________________________

**Student Signature:** _________________________________ Date: _______________

**Parent Signature:** _________________________________ Date: _______________
Appendix I

Teacher Assessment Literacy Training Agenda
Teacher Assessment Literacy Training Agenda

One month before student Assessment Literacy Training

1. Phone call with Scholastic trainer to discuss training needs.
2. Email communication to receive links to training material:
   http://teacher.scholastic.com/math-assessment/scholastic-math-inventory/,
3. Study training material.

One week before students receive Assessment Literacy Training

4. Attempt to give practice SMI assessments to a lower grade level.
5. Train treatment group on Assessment Literacy (see Appendix J).
Appendix J

Student Assessment Literacy Training Agenda
Student Assessment Literacy Training Agenda

Day 1:

- **10 minutes**
  - Attend to daily classroom duties such as, take roll, answer questions from previous day’s work, and correct any homework assignments.

- **35 minutes**
  - Explain objective of Assessment Literacy Training
    - Objective: to gain the assessment literacy skills needed to improve success on summative assessments.
  - Define assessment literacy
    - The understanding teachers and students have to use data provided by formative assessments to adjust learning experiences to gain the required knowledge for success on summative assessments.
  - Give examples and non-examples of feedback on a paper-and-pencil quiz.
  - Example of feedback:
Non-example of feedback:

Homework: give each student a completed sample quiz from a 6th grade math class. Each student is to write useful feedback on the quiz. These will be discussed in small groups the following day.

- 5 minutes
  - Wrap up and end class.

Day 2:

- 5 minutes
  - Attend to daily classroom duties.

- 40 minutes
  - Small group share of feedback
  - Whole group share of feedback
  - Explain the purpose of SMI
    - To assess for student preparedness to learn a new concept. SMI will determine if the student has gaps in their knowledge and give suggestions on what concepts need to be addressed.
  - Show samples of SMI feedback and teach students how to interpret them.
# Student Progress Report

## SMI Test History

<table>
<thead>
<tr>
<th>Test Date</th>
<th>Math Fact Screener</th>
<th>Performance Level</th>
<th>Normative Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>09/10/13</td>
<td>Proficient</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Indicates possible need for work on fact fluency.
- Indicates that student is fluent on tested facts.

For students in Grade 2 and above, performance levels describe the performance of students in Algebra I and Math 1.

### Using This Report

**Purpose:** This report shows a student’s SMI test history. It also provides instructional recommendations based on the test taken in a time period. SMI reports Performance Levels for students in grades 2-12. Other grades are not assigned Performance Levels.

**Follow-Up:** Share information with individual students, noting changes in performance from test to test. Investigate any significant decline in progress. For instructional planning, utilize information provided in instructional recommendations.

---

**Test Date and Score**

- **Grade 5 Year-End Quantile® Range**
- **BMI = Emerging Mathematician**
- **Algebra Readiness Quantile (100IQ)**
- **College and Career Readiness Quantile (1400IQ)**

* Grade Mean is the average Quantile® of all students in this school based on their last test.
Discuss possible places to find resources to address the gaps in knowledge.

- Textbook, online textbook, online resources.

- Homework: give each student a sample SMI report and have them find a resource to fill a gap in knowledge.

- 5 minutes

- Wrap up and end class
Day 3:

- 10 minutes
  - Attend to daily duties
  - Whole class discussion of resources suggested from previous day’s assignment.

- 40 minutes
  - Take first SMI benchmark assessment.
Appendix K

Sample of Online Version of IMMS Survey
1. When I first looked at this unit, I had the impression that it would be easy for me. *

1 2 3 4 5
Not true 0 0 0 0 Very true

2. There was something interesting at the beginning of this unit that got my attention. *

1 2 3 4 5
Not true 0 0 0 0 Very true

3. This material was more difficult to understand than I would like for it to be. *

1 2 3 4 5
Not true 0 0 0 0 Very true

4. After reading the introductory information, I felt confident that I knew what I was supposed to learn from this unit. *

1 2 3 4 5
Not true 0 0 0 0 Very true
CURRICULUM VITAE

SHERYL JEAN RUSHTON

Business Address:          Home Address:
Channing Hall
International Baccalaureate Charter School    9070 Sunrise Circle
13515 S. 150 E.    Sandy, UT 84093
Draper, UT 84020    hm: (801) 944-1034
(801) 972-2709    cell: (801) 573-7937
rushtonsj@comcast.net
Email: srushton@channinghall.org

Education

Ph.D.  March 2014
  Education, Utah State University
  Specialization: Curriculum and Instruction;
  Emphasis: Mathematics Education and Leadership
  Dissertation: The Impact of Computer-Adaptive Benchmark Data and Assessment
  Literacy on Student Achievement and Motivation in Mathematics

M.Ed.  December 2004
  Education, Utah State University
  Specialization: Curriculum and Instruction;
  Emphasis: Instructional Technology

B.S.  June 1987
  Education, Utah State University
  Emphasis: Mathematics, Computer Science, Statistics

EMPLOYMENT HISTORY


Responsibilities include teaching mathematics courses on the middle school level
(courses include: Intermediate Algebra, Geometry, Elementary Algebra, Pre-Algebra,
General Mathematics 7); incorporating the new Utah State Mathematics Core (also
known as Common Core State Standards, CCSS); coordinating class schedules;
representing math department in school and state wide functions; develop curriculum;
assist in creating school policy manual; implement International Baccalaureate
philosophies; and coaching State Math contest and Math Count’s teams.
Utah State University: Teacher Education and Leadership (2013 – Present)
Responsibilities include teaching TEAL 4630/6630 Middle Level Math Methods course through broadcast system.

University of Phoenix: Teacher Education (2013 – Present)
Responsibilities include teaching/facilitating MTE/533 Curriculum Constructs and Assessment: Science and Mathematics, an online course.

Salt Lake Community College (1995 – 2007)
Responsibilities include teaching College Algebra, Intermediate Algebra, Elementary Algebra, Pre-Algebra, and Developmental Mathematics (live courses and online courses); maintaining 12-18 credit hours as an adjunct teacher and 20 credit hours as a full-time faculty; participating on Math 1090 book review committee; being a member of Adjunct Faculty Senate Committee; and integrating technology in teaching.

Responsibilities include development and design of math courses and other UDOT specific courses; and teaching Geometry, Elementary Algebra, Pre-Algebra, and General Mathematics courses (traveling to distant learning sites and via “EdNet” - Salt Lake Community College’s Educational Network).

Responsibilities included teachings mathematics courses on the Jr. High level (courses include: Geometry, Elementary Algebra, Pre-Algebra, and General Mathematics); coaching State Math contest and Math Count’s teams.

Responsibilities include teaching mathematics courses on the Jr. High school level (courses include: Intermediate Algebra, Geometry, Elementary Algebra, Pre-Algebra, General Mathematics 7); coordinating class schedules; representing math department in school wide functions; teaching homebound students; and coaching State Math contest and Math Count’s teams.

COLLEGE TEACHING

Utah State University (2013 – present)
School of Teacher Education and Leadership

Course Taught – Utah State University
TEAL 4630/6630 Middle Level Math Methods course
University of Phoenix (2013 – present)
School of Teacher Education

Course Taught – University of Phoenix
MTE/533 Curriculum Constructs and Assessment: Science and Mathematics, an online course.

Salt Lake Community College, Utah (1995 – 2007)
Mathematics Department/ Developmental Mathematics Department

Courses Taught – Salt Lake Community College
MATH 1050 – College Algebra
College Algebra satisfies quantitative literacy requirements for students planning to take calculus. Topics include polynomial, rational, exponential, and logarithmic functions; matrices; conics; sequences and series; and mathematical induction.

MATH 1010 – Intermediate Algebra
Includes linear and quadratic equations; inequities; polynomials; rational expressions; radicals; negative and rational exponents; complex numbers; linear systems; introduction to functions; logarithms; and exponential functions.

MATH 0990 – Elementary Algebra
Includes linear equations, systems, polynomials, factoring, graphing, and inequalities. It also includes rational and radical expressions and equations.

MATH 0950 – Pre-Algebra
Includes integers, linear equations, polynomials, and graphing. It also includes a review of fractions, decimals, and percents.

MATH 0900 – Basic Mathematics
Introduction to basic mathematics, including operations with whole numbers, fractions, decimals, proportions, and percentages.

GRANTS FUNDED

STATE AND LOCAL SERVICE – LEADERSHIP ACTIVITIES

Demonstrated Competency Assessment Writing for Secondary 1, 2, and 3 Committee Member (2013 – 2014).

Invited by the State of Utah to create questions for the Demonstrated Competency Assessment, Secondary Math 1.

Grant Review Committee Member (2013 – 2014).

Utah State Office of Education Numeracy Grant Review Committee.

Invited by the state to review grant letters of intent and award grant money.

District Staff Development Director (2010 – present).

Channing Hall District Staff Development Director.

Assist in the organization of staff development coaching and manipulative use instruction mainly for mathematics.

Committee Member (2010 – present).

State Secondary Math Instructional Material Advisory Committee.

Invited by the state to evaluate instructional material submitted to the by publishers for recommendation to the schools in the state of Utah

Committee Member (2007 – present).

Utah State Mathematics Education Coordinating Committee.

Represent Channing Hall at statewide meetings. Collaborate with math coordinators from several Utah school districts and charter schools in the development of statewide mathematics programs.


Channing Hall District Mathematics Coordinator.

Coordinate and organize the Mathematics program at Channing Hall charter school for grades K-8.

Math Department Chair (2006 – present)

Channing Hall Mathematics Department Chairperson.

Coordinate Mathematics department with school wide functions. Incorporate new mathematics curriculum K-8.

Math Department Chair (1990 – 1993).

Mueller Park Jr. High Mathematics Department Chairperson.

Coordinated class schedules, conducted monthly departmental meetings, represented the math department in school wide functions.
PRESENTATIONS
Professional Presentations


Invited Presentations

Invited presenter for USU GEAR UP grant, with lead Project Investigator - Eric Packenham, at the UCTM conference. (2013, November).

Invited presenter for TEAL 7551: Mathematics Education Research Foundations (for Dr. Patricia Moyer-Packenham) (2013, March).


LEADERSHIP & SERVICE

Reviewer (2011) *Association of Mathematics Teacher Education* (ATME), Reviewed proposals for the annual meeting.

PROFESSIONAL SERVICE

Search Committees

- Search Committee Member: Assistant Head of School for Channing Hall. (2009 – 2010).

Committee Memberships

- Middle School Task Force: Meets bi-weekly to provide and gather feedback on students and actions taken with students to improve their educational experience (2006 – present).
• Channing Hall Policy Manual Development Committee: Create the policies for a newly developed charter school (2006 – present).
• Land-Trust committee: Meet monthly to determine the needs of the school and prepare applications to receive Utah Land-Trust money (2007 – 2010).
• Salt Lake Community College Adjunct Faculty Senate Committee: Meet monthly to discuss issues and make improvement for the Adjunct Faculty Members (2005 – 2006).

Curriculum Development
• Committee chair for Mathematics curriculum review and adoption board (2012 – present).
• Develop cross curricular lessons in conjunction with Science, Language Arts, Humanities, Art, and CTE departments (2006-present).
• Developed the curriculum for Channing Hall Middle School Mathematics department (2006-present).
• Curriculum Writing of 7th grade Mathematics for Utah Common Core State Standards (2011).
• Curriculum Writing of 8th grade Mathematics for Utah Common Core State Standards (2011).
• Curriculum Writing of Secondary I Mathematics for Utah Common Core State Standards (2010).
• Designed Blackboard courses for UDOT: Transportation Technician Training (Salt Lake Community College) surveying courses (2005).
• Salt Lake Community College Math 1090 Book Review for new curriculum (2005).
• Developed the curriculum for the Pre-Algebra and Elementary Algebra courses for UDOT: Transportation Technician Training (Salt Lake Community College Instructional Technology) (1997).

Other School Service

Awards & Professional Recognition
• Nominated for Presidential Award for Teachers (1989).
• Elected a member of Who’s Who Among America’s Teachers

Professional Affiliations & Leadership Roles

NATIONAL COUNCIL OF SUPERVISORS OF MATHEMATICS (NCSM)
• Member (2012-present).
NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS (NCTM)
- Member (2011-present).
- Committee member (1986).

UTAH COUNCIL OF TEACHERS OF MATHEMATICS (UCTM)
- Member (2006-present).