12-2013

The Effects of Standards-Based Grading on Student Performance in Algebra 2

Rachel Beth Rosales
Western Kentucky University, rachel.rosales@daviess.kyschools.us

THE EFFECTS OF STANDARDS-BASED GRADING ON STUDENT PERFORMANCE IN ALGEBRA 2

A Dissertation
Presented to
The Faculty of the Educational Leadership Doctoral Program
Western Kentucky University
Bowling Green, Kentucky

In Partial Fulfillment
Of the Requirements for the Degree
Doctor of Education

By
Rachel Beth Rosales

December 2013

THE EFFECTS OF STANDARDS-BASED GRADING ON STUDENT PERFORMANCE IN ALGEBRA 2

Date Recommended 10-29-2013

Marge Maxwell

Marge Maxwell, Director of Dissertation

Rebecca Stobaugh

Rebecca Stobaugh

Janet Tassell

Janet Tassell

Carl A Fox 11-22-13

Dean, Graduate Studies and Research Date

This work is dedicated to my family, for without their constant support I would have not finished.

To my husband Ben, who had to endure a considerable amount of take-out and frozen pizzas for supper, laundry piled to the ceiling, numerous weekends alone while I worked, and an inordinate amount of crying and yelling at times.

To my daughter Morgan, who didn't flinch when I told her I was going to college at the same time and place as she was.

To my parents John and Judy, who have always inspired me to work hard and never quit, who have helped me understand the value of the pursuit of knowledge.

And to my late mother-in-law Dorothy, who was always a cheerleader and supporter of higher education. I wish she could be here to celebrate this accomplishment with me.

ACKNOWLEDGEMENTS

Although this project bears my name, numerous people along the way have supported, helped, or encouraged me and without whom I would not have finished. Dr. Matthew Constant, pioneer doctoral student and trailblazer with Cohort I, was most inspirational in convincing me to even start this endeavor. All along the way he has provided advice, support, guidance, suggestions, and a soft ear when one was needed. He has read and critiqued many applications and papers along the way and has simply been my friend through it all.

My colleagues at Apollo High School have bent over backwards to help me make it through. They have covered my classes at times, endured tales of woe and despair, and celebrated at the milestones along the way. The administration has been understanding and cooperative when a day off was needed or when lesson plans weren't entirely perfect. I would be remiss if I didn't thank my students during this time who went along with me, and always gave me their very best efforts, even when they didn't quite understand what was happening.

There is absolutely no way this project could have been undertaken without the support and assistance of a dear friend and colleague, Becky Mann. She offered her advice, worked through the methodology with me, helped me with issues as they arose, and offered her classes as guinea pigs for the study. I am forever indebted to her.

Dr. Marge Maxwell has been a constant source of inspiration, patience, wisdom, and assistance from the very beginning of this project. She has endured numerous topic changes and uncertainty from me and deserves an award for the way she has handled me.

I'm sure she felt frustrated and tired, yet she always was willing to listen to any crazy notion I had and has worked tirelessly to help me present a complete and final product.

The other members of my committee, Dr. Becky Stobaugh and Dr. Janet Tassell, have been sources of encouragement and advice throughout the design and implementation of this project. They both shared in the vision of what I wanted to do and encouraged me along the entire way.

Finally, my cohort, Cohort V, provided kindred spirits sharing this journey with me. Each member of that cohort touched me or helped me in some special or unique way, whether it was a simple pat on the back or sharing in a laugh over a quick lunch during a stats class. Without each member of this group, we would not have been a complete cohort. There are two members with whom I developed strong ties and will be forever bonded in friendship. David Oliver and Dana Clayton, providers of good coffee and pumpkin spice donuts, always will have a special place in my heart.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

THE EFFECTS OF STANDARDS-BASED GRADING ON STUDENT PERFORMANCE IN ALGEBRA 2

Rachel Rosales December 2013 89 pages

Directed by: Marge Maxwell, Rebecca Stobaugh, and Janet Tassell

Educational Leadership Doctoral Program Western Kentucky University

The use of standards-based grading in American public schools is increasing, offering students, parents, and teachers a new way of measuring and communicating about student achievement and performance. Parents indicate an appreciation for this method of grading, and students at the elementary grades (K-6) have improved standardized test scores in reading and math as a result of its implementation.

This study seeks to determine whether standards-based grading has the same effect on students at the high school level (grades 9-12) by comparing end-of-course test scores and posttest scores of Algebra 2 students enrolled in a standards-based graded classroom with to those enrolled in a traditionally-graded classroom. Two teachers each taught two classes of Algebra 2 and graded one class using standards-based grading and one class using traditional grading methods. Students at both the honors level and the regular level of mathematics were included in the study.

Honors students performed better than regular students on both assessments, but no significant difference was found between the performance of traditionally-graded students and the students who were graded with standards-based grading.

The results of this study indicate that standards-based grading may offer improved methods of communication between teachers, parents, and students and may give students a new perception of learning. Standards-based grading strategies require careful planning, dedication, and follow through. It is not an endeavor to be entered into lightly,

but rather, the appropriate amount of time, resources, and preparation can provide students the chance to truly learn content at a mastery level. The mastery of material will, in turn, translate into higher success on high-stakes testing.

CHAPTER I: INTRODUCTION

Grades have been a deep, mysterious secret language, understood by only those who maintain them, yet used widely by parents, universities, scholarship committees and students as a sure indicator of what a student knows and can do. It has been only within the last decade that the foundation of the all-powerful grade has been questioned. In addition, student performance has been placed under a microscope and observed by stakeholders at all levels, from classroom teachers to the President of the United States.

In 2002, President Bush signed into law the No Child Left Behind Act, placing education on a course for reform. In 2009, President Obama issued a call to all Americans to complete at least one year of post-secondary studies, and his Race to the Top legislation renewed the drive for further educational reform (Strauss, 2012). Students in American schools are performing at a much lower achievement level than their counterparts in other countries (Duncan, 2012), and state education leaders have been under pressure to show evidence of improved student performance. A significant component of these reform efforts has been the creation of standards in English and mathematics, with 45 states and 3 United States territories adopting the Common Core Standards (National Governors' Association Center for Best Practices & Council of Chief State School Officers, 2012). For the first time in American history there is a sense of a national curriculum and an urgency to reach a level of proficiency. Across the nation public school educators have searched for new and better teaching methodologies, assessment methods, and instructional techniques to address the needs of a diverse student population in meeting these expectations.

The creation and adoption of Common Core Standards (CCS) have caused many teachers to pause and consider the degree to which their instruction is aligned to these standards. For many, the thrust toward heightened alignment has provided an impetus to deliberate the instruction presented to students, the techniques of assessment used to determine student achievement, and the communication of student performance and growth.

As a result of the intense effort to reform education, school districts have realized they are in need of dynamic and profound changes. The culture, which has allowed education to thrive during previously stable times, is not the same culture that is necessary for navigating the current tumultuous era of reform. Altering a culture is not an easy task; while raising expectations is a priority, doing so is not as simple as presenting new standards and accountability models. Black and Wiliam (1998) described the problem succinctly. They reported that current policies, both domestic and foreign, approach education as a "black box" (p. 2) into which inputs (teachers, standards, rules, policies, exams, etc.) are fed (see Figure 1). The expectation is that a prescribed set of outcomes will follow (higher scores, smarter students, satisfied teachers, etc.). However, little directive is given to the teachers, and no one is considering the events that occur within the box.

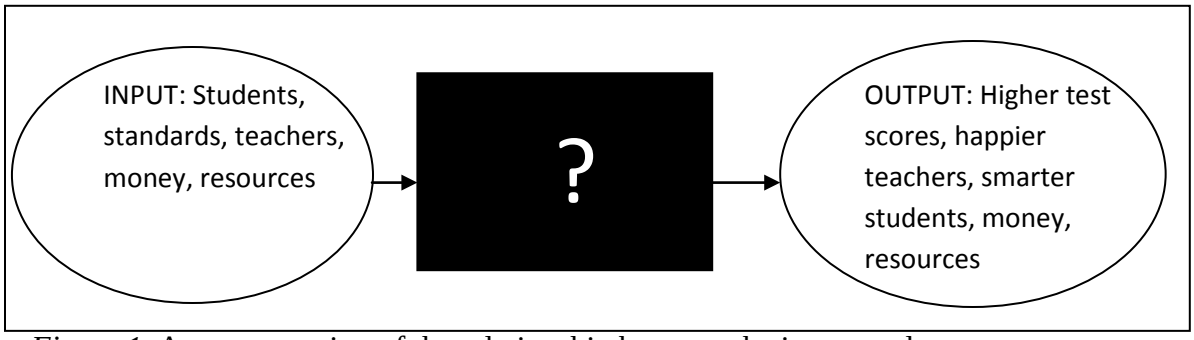

Figure 1. A representation of the relationship between the inputs and outputs of a classroom, based on the ideas of Black and Wiliam (1998)

The mandate is clear. Educators are to raise their expectations and provide the support necessary for students to advance farther. Schools that fail to provide evidence of growth in achievement face serious consequences (O'Shea, 2005). The intense focus on higher standards requires a tighter alignment to grading mechanisms. To this end teachers and school districts are considering the implementation of standards-based grading (SBG).

Theoretical Basis for the Study

A standards-based approach to teaching and learning is based on the theory of formative assessment. Sadler (1989) outlined a theory that details the function of formative assessment in the growth and development of student expertise. According to his model, three key factors are necessary for an assessment strategy to assist in moving a learner forward in his performance ability. First, the learner must understand the desired outcome toward which to work. Second, the learner must be able to compare the current level of performance to that which is desired; and, third, some action must exist that can be employed to help bridge the gap between the two (Sadler, 1989). Standards-based grading seeks to embody these three facets, resulting in an assessment structure designed to move learners forward and increase their skill set.

While a general theory for grading has yet to be explicitly defined (Brookhart, 1994), a framework has been presented which attempts to identify the role of assessment in motivating effort and achievement (Brookhart, 1997). Within this theory, Brookhart contended that student perceptions of the assessment task and of their own self-efficacy affect the amount of effort invested in the event. If the task is perceived as valuable and

within the abilities of the student, that student will exert an appropriate amount of effort and will experience achievement. Standards-based grading can define what is necessary for success, making the task valuable to the student, and offer practical methods for personal growth and development, making the task feasible.

Significance of the Problem

Grades serve a variety of purposes; historically, grades have been used to motivate, sort, rank, and qualify students for college entry and scholarships (Brookhart, 2011; Guskey, 2001). Explicitly defining a set of standards, refocusing the goals of the school, and expecting to attend post secondary education have resulted in a shift in the purpose served by grades. Rather than ranking students, grades need to inform them of their progress and areas of need. Rather than sorting students, grades need to provide honest feedback to parents. Grading methods need to distinguish between performance and non-academic indicators, such as effort. Students need to know where they stand relative to the competencies of the course and need to have an accurate notion of how they will perform on high-stakes tests.

For years, teachers have determined how grades in their classrooms have been calculated, often doing so in isolation, without input or collaboration from other teachers; the consequences of teacher-determined grading methods have been numerous. Non-academic attributes such as completing assignments on time, neatness of assignments, effort, and cooperation have caused an inflation of grades, leaving students with the belief they are achieving at higher levels of performance than they actually are (Godfrey, 2011). In the 1990s, two ACT researchers found that student GPAs were outpacing their ACT scores. The average GPA of all high school graduates in 1990 was 2.68; by 2000, it

had increased to 2.94. The College Board found an increase of 5% from 1994 to 2004 in the number of students reporting a high school grade average of an A, but the scores of those same students dropped in both sections of the SAT (Zirkel, 2007). Students appear to be doing better in school, but they are performing to a lesser degree on standardized exams.

Inequitable grading is another concern for educators and students. Two instructors teaching the same course each may have different opinions regarding what constitutes an acceptable level of mastery, as well as the weight of a variety of additional factors in the final grade. For example, homework may constitute a considerable portion of a final score in some classes, while in other classes it may be receive minimal weight. Students doing the same quality of work may receive different grades, based on who taught the course (McMillan, 2001).

The effort to reform education cannot exist without a concerted focus on standards and grading. Two issues must be confronted if student achievement is to be accurately measured; one is preparation for end-of-course exams and other standardized assessments and the second is the discrepancy between report card grades and high-stakes testing performance. Standards-based grading (SBG) provides one possible system for addressing both of these issues by helping teachers align classroom marks to external assessments. SBG has changed the way teachers communicate grades on report cards (Tierney, Simon, & Charland, 2011) and has resulted in improved classroom grades (Guskey, Jung, & Swan, 2011).

Improved performance on teacher-made exams provides a good start to raising achievement levels. However, if students fail to exhibit the same improvement on

externally assessed instruments, school systems will not make advances toward meeting the mandates from the government. If SBG is to be considered a worthwhile endeavor, it must be shown not only to increase classroom grades, but also to improve scores on state and national exams. This study will look at the effects of SBG on student performance on end-of-course assessments in Algebra 2.

Statement of the Problem

Standards-based grading is a growing topic of conversation across the United States, with numerous teachers adopting it as daily practice in the classroom; the problem is that the effectiveness of SBG to raise student performance on external assessments has not been measured at the secondary level. This study will address the effectiveness of standards-based grading as a mechanism for improving student performance on summative assessments in Algebra 2.

Rationale for the Study

While standards-based grading seems an appropriate means of addressing grade inflation and equitability issues, the question of its ability to improve student performance on external assessments still needs to be researched. A few studies have highlighted the effectiveness of SBG to improve student performance in the classroom and on the positive parental and teacher perceptions of SBG (Craig, 2012; Guskey et al., 2011; Haptonstall, 2010; Hardegree, 2012). However, few studies have addressed the ability of SBG to impact student performance on high-stakes or externally written tests. The emphasis on standards and accountability requires teachers to use measures of proven worth. If SBG is to become the norm of grading and assessing, educators need a solid basis of evidence to support its implementation.

Research Questions

The central research question that guided this study is: Does standards-based grading affect student performance in Algebra 2? Under this central research question, the specific research questions addressed in this study include:

1. How do the grading method and the student achievement level impact the Algebra 2 end-of-course assessment?
    a. How does the grading method (standards-based grading and traditional grading) impact the results of the Algebra 2 end-of-course assessment?
    b. How does the student achievement level (honors and regular) impact the results of the Algebra 2 end-of-course assessment?
    c. Is there interaction in student performance on the Algebra 2 end-of-course assessment between the grading method and the student achievement level?
2. How do the grading method and the student achievement level impact the classroom posttest results?
    a. How does the grading method (standards-based grading and traditional grading) impact the classroom posttest results?
    b. How does the student achievement level (honors and regular) impact the classroom posttest results?
    c. Is there an interaction in student performance on the classroom posttest results between the grading method and the student achievement level?

Definitions and Terminology

For the purposes of this study, the following terms and definitions were used:

Achievement level — used as an independent variable and defined by the level of math class the students were taking during the study, honors or regular.

End-of-course assessment (EOC) — the state mandated test given at the completion of Algebra 2, taken by all students enrolled in the course. The EOC exam consists of two 38-question multiple-choice sections. Students had 45 minutes to complete each section.

Grading method — used as an independent variable and defined by the method of grading implemented in the four classes, standards-based grading or traditional grading.

Learning target — using student-friendly language, learning targets written by the classroom teacher based on the curriculum standards, that described the skills, processes, and information students developed during the course of instruction. Learning targets are more than objectives; they guide the learning process. (Moss & Brookhart, 2012).

Standards-based grading (SBG) — method of grading students, based on performance on a set of defined objectives and learning standards (Guskey et al., 2011). In SBG, students were individually graded on each learning target and given opportunities to retest in areas of deficiency.

Traditional grading — method of grading students based on percentages of correct responses. For this comparison, traditional grading indicated students received marks for homework assignments weighted at 10% of the final course grade and marks for tests and quizzes which were weighted at 90% of the final course grade. Students in the traditionally-graded course were not required to re-test any areas of deficiency, although they may have been permitted to do so on a case-by-case basis.

Summary

This chapter has presented a brief introduction of educational reform and the concept of standards-based grading, while establishing the need for further research into the effects of SBG on student performance. A brief history of standards-based education has been presented, providing the background necessary to understanding the context for the study. Chapter II will look at the literature relevant to the study and will further build a foundational basis for a study of SBG in a classroom setting. Chapter III will present the details of the classrooms involved in the pilot program and the methodology used to guide the research. Chapter IV will provide the statistical results of the tests and Chapter V will offer a synopsis of those results with interpretations and implications for future research.

CHAPTER II: REVIEW OF THE LITERATURE

Introduction

The high stakes prevalent in education raise the accountability level of teachers in all aspects of the classroom. Quality of instruction and adherence to the standards are two elements for which teachers are held to increasingly higher expectations. In addition, accurate grading practices need to be implemented so that all stakeholders have an accurate picture of what students know and can do. The increased focus on standards makes it important to apply a grading system that provides students and parents with a true indication of ability level. The inclusion in grade calculations of non-academic attributes, sometimes referred to as soft skills (i.e., timeliness, neatness, effort, and cooperation), have caused, in some cases, an inflation of grades, leaving anyone who looks at a report card to believe that students have reached a higher level of achievement than they actually have (Godfrey, 2011). Students' knowledge also can be obfuscated negatively when zeros are used as punishment for not displaying the skills listed above. Penalizing students with zeros for misbehavior, failing to meet a deadline, or sloppy work can result in a lower grade than what the student deserves (Guskey, 2004). True content knowledge can be masked when these soft skills are given significant weight in a student's final grade.

To address the concerns of accurate and valid grading, some teachers and school districts have implemented standards-based grading (SBG). While teachers and parents appear to be pleased with the results (Guskey et al., 2011), the question remains as to whether the students' achievement has been positively affected. The purpose of this study is to compare the performance of Algebra 2 students in traditionally-graded courses to the

performance of students in SBG courses to determine whether SBG has had an impact on exam scores.

Theoretical Framework

While no universally accepted definition of standards-based grading exists, key elements can be found that are common to most interpretations of the concept; among these are the establishment of student-friendly objectives and goals, student reflection and self-assessment, opportunities for demonstrating mastery through assessment, and effective feedback to the student from the teacher. Each of these components is critical to the success of SBG and is grounded in the theory of formative assessment.

The theory of formative assessment is a relative newcomer to the educational arena. Black and Wiliam (2006) outlined a theory after reviewing findings from the King's-Medway-Oxfordshire Formative Assessment Project (KMOFAP) and the Berkeley Evaluation and Assessment Research project (BEAR), programs designed specifically to improve teaching and learning through the use of formative assessment. By carefully analyzing classrooms that had implemented formative assessment techniques and noting the specific elements which had undergone change, the researchers were able to collect and classify a vast and diverse set of issues into a framework to provide a basis for further exploration and study. Their study yielded a set of four elements which, in combination, constitute a theory of formative assessment.

Formative Assessment

One observed change was in the relationship between the discipline and the teacher's role. In order for students to improve, teachers realized the necessity of achieving quality in both the questions they asked and in the responses given to students.

Attainment of a high quality of interaction between the students and the teachers required teachers to become more attuned to the structure of their particular discipline and to carefully craft questions and responses in ways that would maximize the learning experience for the students. The relationship between the content of the course and the role of the teacher was determined to be very discipline-specific; and, while the manner with which teachers interacted with students varied greatly from one subject to the next, the quality of the interactions did not (Black & Wiliam, 2006).

Second, an adjustment was noted in the teachers' view of their role in the learning process. Prior to the project teachers tended to plan classes from an activity perspective, with a focus on what the students would do during the class; after participating in the project teachers reported a shift to a focus on what the students would learn. This shift resulted in a greater emphasis on creating lessons embedded with "teachable moments" (p. 87), opportunities to provide skillfully crafted feedback to the students. This shift placed the responsibility for learning onto the student and left the teacher with the responsibility to provide quality learning environments (Black & Wiliam, 2006).

Third, feedback transformed the interactions between the students and the teachers. This altered interaction holds a crucial spot in the formative assessment process, providing the key portal through which learners manage to reduce gaps in knowledge. Effective feedback offers learners a mechanism for self-assessing their performance and working to achieve the learning goals. No simple, straightforward recipe exists for giving effective feedback. In addition to consideration for the contexts of the classroom and the structures defined by the subject matter, teachers must understand how students receive

messages as well as how they think in the classroom to be able to offer feedback that will be useful to students (Black & Wiliam, 2006).

Finally, in each classroom of the project, the student's role was changed; students no longer operated passively, waiting for the teacher to feed them information, but instead, became active learners who assumed some responsibility for their own learning. Students also indicated that they felt as if their teachers believed in more than mere test results; they believed in the students. The focus for the students seemed to shift from correct answers on the test to true understanding and achievement (Black & Wiliam, 2006).

Classroom as an activity system. Four changes were observed in the formative assessment classrooms: the teacher's view of her subject, the teacher's view of her role in the learning process, the feedback between students and teachers, and students' views of their own role in learning. These changes allow the subject area classroom to be viewed from an activity system, a "complex formation in which equilibrium is an exception and tensions, disturbances, and local innovations are the rule and the engine for change" (Black & Wiliam, 2006, p. 83, quoting Salomon, 1993, 8-9).

The central elements of an activity system include the individual or subgroup, an object which is acted upon to produce outcomes, tools, the community who share the same objective, division of labor, and the rules. The actions taken to achieve the goals, and the social conditions surrounding the desired goals, create an activity system (Black & Wiliam, 2006). These elements can be connected, as displayed in Figure 2.

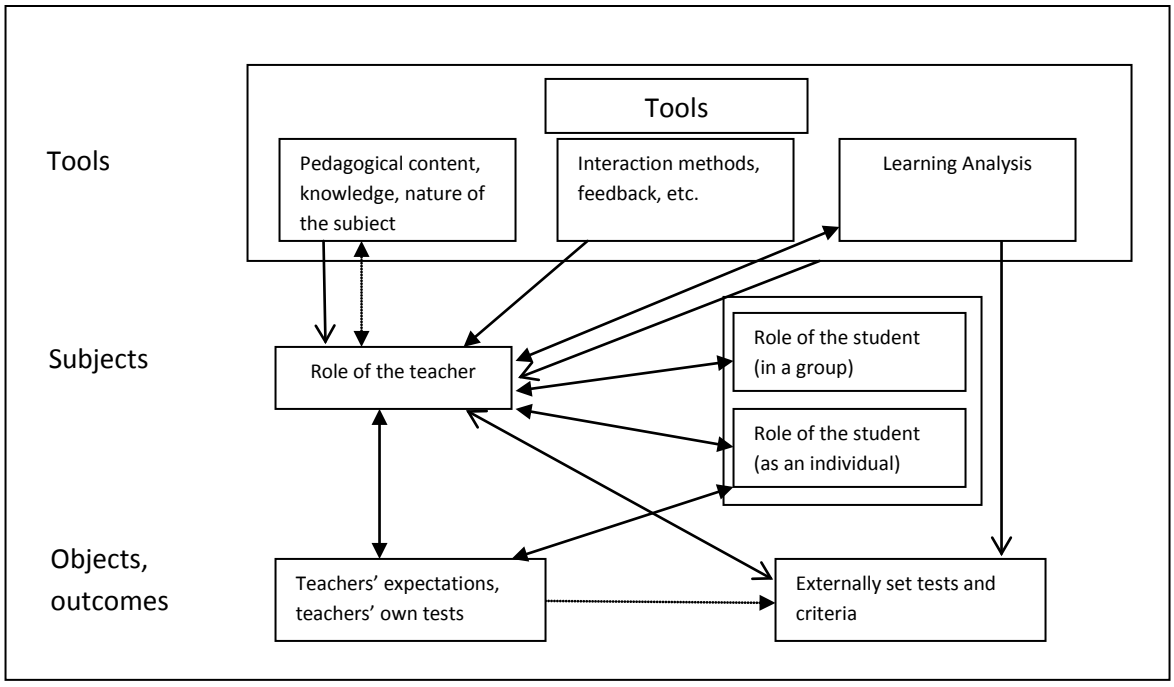

Figure 2. Patterns discovered in the KMOFAP and BEAR projects. Small arrows represent findings from BEAR, large arrows represent findings from KMOFAP (Black & Wiliam, 2006, p. 95).

From this diagram, the theoretical framework for the formative assessment classroom is evident. The tools available (such as quality questions and feedback) alter the role of the teacher, which directly affects the role of the students. Together, these tools and subjects have a direct effect on the expectations set by the teacher and ultimately on externally assessed components such as state tests.

Elements of a standards-based classroom. From this framework, the basic elements of an effective standards-based classroom begin to emerge. First, the learners must know what is expected of them; this can be accomplished through clear, descriptive statements and presenting examples of high quality work to the learners (Sadler, 1989). Black and Wiliam (2009) described this portion of the process as "establishing where they [the learners] are going" (p. 7) and Hattie (2009) reported that articulated outcomes

have a high effect on student performance. Learning goals provide a solid foundation for communication about the skills and concepts a student does or does not know and offers both the teacher and the student clear direction for each unit of study.

After learning goals are clearly understood by students, they must have the opportunity to gauge the gap between their current abilities and the expectations set forth in the learning targets (Sadler, 1989). Stated more simply, students need to be able to ascertain where they currently stand in the learning (Black & Wiliam, 2009). This self-reflection offers students the opportunities to not only develop in their content knowledge, but also develop in their abilities to see how they learn (Marzano, 2006). Self-assessment allows students to participate in two processes important in developing thinking skills: understanding the nature of the task and possessing a conscious awareness of one's own thinking (Gredler, 2011).

Finally, after the students understand the goals and can assess their current gap in skill, there must exist a practical, reasonable action plan that students can undertake to will help them move forward in their learning and close the gap (Sadler, 1989). This idea is evident in the SBG component of reassessment until mastery. Students who learn content and operations may not fully develop their understanding at the same pace as their peers; SBG offers students the option to re-test concepts after further practice, and review provides them time to develop and demonstrate that development at their pace, when they are ready. Feedback plays a significant role at this stage of formative assessment, alerting the student to areas of excellence as well as areas in which growth is still necessary. Good feedback is one of the most powerful tools a teacher can use to influence student performance (Hattie, 2009). Students must be given informative

feedback, letting each one know their own strengths and weaknesses on each learning goal. Without clear, informative communication, SBG cannot be successful. Students and teachers are collaborators, working together to develop the knowledge and abilities of the students.

Grading

Standards-based grading brings assessment to the forefront of the classroom and requires buy-in from the students. Their perceptions of the value and importance of a task are critical elements in the amount of motivation they will exhibit (Brookhart, 1997). Ideally, students would be engaged in tasks they find meaningful, such as writing a letter to someone. However, there are cases in which the meaning must be derived from the importance of the activity to the central understanding of a key concept. For example, students may not find meaning in the process of solving a quadratic equation, but they can understand the importance of that skill to successfully complete a unit in algebra. Brookhart's (1997) proposed framework of assessment maintains that this perception of importance will increase student motivation and cause them to exert the effort required to achieve success. Mastery learning leaves students with positive feelings about class and beliefs that their efforts will yield success (Brookhart, 1997). Classroom efforts that provide an effective mastery-learning setting include sharing of authority, thoughtful task design, and evaluation geared toward progress.

Standards-based grading operates within this framework, establishing appropriate learning targets, and then creating assessment tasks from those targets. Students are given opportunities to build their self-efficacy through practice and the chance to reassess until they reach a satisfactory level of accomplishment. Ideally, a "snowball" effect will occur,

with students experiencing success, thereby building their confidence level, and in turn helping them be more successful.

Origins of Standards-based Education in the United States

The earliest traces of educational standards go back to Ralph W. Tyler in 1934. He argued that objectives and aims should not be eloquent and grand, but instead should be stated in terms that clarify the desired behaviors and outcome of the students (Stones, 2012). In the mid-1950s, Bloom published his Taxonomy of Educational Objectives (1956) to help "specify objectives so that it becomes easier to plan learning experiences and prepare evaluation devices" (p. 2).

Educators have been using standards of learning for decades; however, the concept of an entire educational system focused on a set of prescribed standards is considerably newer. In the late 1970s and early 1980s, policymakers and government leaders began to take notice of the declining educational situation in the United States. Based on international assessments, it was becoming clear that American students were beginning to lag behind their counterparts in other countries. The Secretary of Education at that time, T. H. Bell, created a task force to ascertain the quality of American education; the result of their effort was a report titled *A Nation at Risk* (1983). This report launched public education into a new age of accountability, standards, and assessment (Hamilton, Stecher, & Yuan, 2008).

A Nation at Risk (1983) contained a list of concerns regarding student achievement. At that time, approximately 23 million Americans were classified as functionally illiterate, 13% of whom were 17 year olds; students were scoring lower on standardized achievements than they were scoring when Sputnik was launched 26 years

earlier; the number of remedial math courses was increasing; and the military was spending millions of dollars on training programs in basic skills. Additionally, the report issued a list of findings regarding the curriculum offered in schools at that time. According to the task force, high schools had diluted the content offered to students to the point that there was little focus or purpose. Students were free to choose their courses, and few were selecting rigorous and challenging paths of study. Twenty-five percent of the credits earned by students on a general track were in physical education, remedial courses, or work experiences. Only 31% of graduates completed intermediate algebra; in schools where calculus was offered, it was completed by only 6% of the students. Expectations in schools were at a minimal level, and students were ineffectively spending time in and out of the classroom (National Commission on Excellence in Education, 1983).

The task force continued with recommendations for improving the quality of education. In essence, this document published the first official list of standards to be implemented. English, mathematics, science, social studies, and computer science were addressed as basic elements of a strong curriculum, with each subject receiving a specific list of suggested course offerings. The report also explicitly called for “more rigorous and measureable standards” (Recommendation B: Standards and Expectations) as well as the administration of achievement tests (National Commission, 1983). From this beginning, standards-based education found its roots and emerged as a key factor in school reform. Following the release of *A Nation at Risk,* state governments responded by creating their own task forces designed to scrutinize their own systems of education and develop action plans for improvement. Although some states experienced success in their reform efforts,

the nation as a whole did not meet the expectations set forth by the commission (Bell, 1993).

The Federal Government

The role of the federal government in the standards-based reform initiative can be credited to President George H. W. Bush's Education Summit in 1989. During this meeting between President Bush and the state governors, the call was made for the government to develop strategies for improving education to prevent American students from continuing to fall behind their international peers (Hamilton et al., 2008). Although the goals developed at this summit meeting never became official law, they marked the first time education improvement became a topic of national political interest.

In 1994, President Bill Clinton signed into legislation the Goals 2000: Educate America Act, which reinvigorated the notion of improving education by creating standards of achievement for all students. Among the goals was the expectation that all students would "leave grades 4, 8, and 12 having demonstrated competency over challenging subject matter including English, mathematics, science, foreign languages, civics and government, economics, arts, history, and geography" (sect. 102 3A), while increasing the graduation rate to 90% (HR. 1804, 1994). Goals 2000 made it necessary to increase the assessments students were required to take so that the degree to which these goals were being met could be measured. In addition, it started the process of identifying exactly what information students should know and what skills they should be able to perform on these assessments.

In 1994, Congress reauthorized the Elementary and Secondary Education Act, which required each state to develop standards for all students in kindergarten through

twelfth grade and introduced adequate yearly progress. States were required to demonstrate evidence of growth and improvement, but this reauthorization did not provide any deadlines for completion; therefore, no sense of urgency was present. Additionally, no consequences occurred if no evidence of growth or improvement was found. By 1997, only 17 states had succeeded in defining specific standards in core content areas (Rudalevige, 2003).

President George W. Bush, Jr., working with the legislation and ideas already in place in education, formalized the No Child Left Behind Act (NCLB) in 2001. NCLB mandated yearly testing and required states to release annual report cards showing their performance. Student achievement was broken down into sub-populations, and failure to meet the needs of each group of students meant schools did not meet the requirement of adequate yearly progress (Rudalevige, 2003). Expecting all students to reach proficiency meant states and schools could not ignore the mandate for improvement, and standards-based education became the norm for educators.

Common Core Standards

In 2010, the Common Core Standards (CCS) in English/language arts and mathematics were released and presented to the states jointly by the National Governor's Association (NGA) and the Council of Chief State School Officers (CCSSO). Although adoption of the standards remained at the discretion of each individual state department of education, the ante to do so was increased when President Obama announced the Race to the Top Assessment Program. Any state wishing to be in the competition to receive millions of dollars was required to adopt the CCS. In addition to adopting the CCS, states were required to develop "assessments that are valid, support and inform instruction,

provide accurate information about what students know and can do, measure student achievement against standards" (U.S. Department of Education, 2012).

A Nation at Risk (1983) ushered education into a new era; the 25 years since have seen renewed purposes and new directives for all students. To be educated in America no longer means simply showing up for class and putting in seat time. Students are held to new, higher levels of accountability, and teachers are accountable for making the changes necessary to support that expectation. Methods, assessments, techniques, and grading must align to the standards, point students toward the goals, and communicate effectively to parents how their children are performing in light of the standards.

Evolution of the Purpose of Grades

Grading can be traced back to 413 B.C., when survivors of the Athenian army were thrown into quarries; their lives or release from capture depended on their ability to quote verses of Euripides (Cureton, 1971). In this early beginning, grading simply meant passing or failing. Creativity, neatness, and effort had no role in this test. The first achievement test occurred in the sixth century in China with the administration of the first civil service examination consisting of writing, reciting classic works from memory, and analyzing political problems (Crozier, 2002).

The earliest forms of student grading placed the subjectivity on the shoulders of the teacher. Grading meant organizing student work into categories of deficient, medium, or high; the concept of grading was merely the sorting and ranking of students (Cureton, 1971). In the latter half of the nineteenth century, teachers began using percentage scores to report student achievement. In 1879, Josiah G. Fitch, a school inspector in England, presented a lecture at Cambridge University regarding a grading method whereby

teachers could judge the achievement of the students. Contained in his outline was a suggestion to reserve 10% of the marks for items such as “style, neatness and finish, and general skill” (Cureton, 1971, p. 3). Additionally, he suggested that, after tabulating the marks earned, the teacher should review the piece as a whole to determine whether the total genuinely represented the merit of the whole work. If not, he suggested compensating by adding marks for general ability. The subjectivity and judgment of the teacher were very evident in the awarding of final grades. It is interesting to note that Sir Fitch also believed it would be unlikely for students to earn the maximum number of marks possible; in his opinion, a good paper would receive about 75% of the marks, and a passable paper would receive about 50% (Cureton, 1971).

Not until the beginning of the twentieth century did educators begin to search for a more scientific method of grading; most institutions used the percentage method, but the requirement for passing varied from institution to institution. In some cases, students needed a 60% to pass, while in other cases a 75% was required. Additionally, studies as early as 1912 found that teachers did not grade consistently; considerable variability was found in the marks given by different teachers on the same assessments. From these analyses, the conclusion was drawn that, while teachers could score papers reliably if limited to percentages that were multiples of five, such as 90 or 95, they could not score reliably when using more precise grades such as 92, 93, or 94. As a result, many educational institutions adopted a letter grading scale, issuing an A for a band of percentages such as 90-100 (Cureton, 1971).

Grading Practices

More recent surveys of teachers reveal that the same purposes of grading are still prevalent today. Many of these surveys indicate that teachers believe grades should be used to accurately indicate a student's academic level of accomplishment, and a concern surfaced regarding a lack of consistency from teacher to teacher. Others believe the grade a student receives should include an element of effort, while others believe grades motivate students to perform better (Brookhart, 2011). Formal training in grading practices is virtually non-existent, leaving inexperienced teachers to develop a practice they deem appropriate. Without knowledge of the effectiveness of various strategies, many resort to the only practice they know, the practice they experienced in their own coursework (Guskey, 2004). However, the increase in standards-based educational efforts and initiatives has sparked new thoughts and ideas about how grades should be assigned and reported, and a new method of grading has emerged. Today, two general practices of grading have been adopted by teachers and schools. Many still use a traditional idea of grading, but others are implementing a practice now known as standards-based grading.

Traditional Grading

Traditional grading has its roots in the history of education. It has been in place for centuries, receiving only small tweaks along the way. Most of the U.S. population understand traditional grading methods because those methods have been used since before the twentieth century. Parents are familiar and comfortable with their child receiving a final letter grade for a course and understand the implied meaning of a grade such as a "B" or 94% (Guskey, 2011).

While some of the small details of traditional grading vary from teacher to teacher, the basic principle is that students receive a numeric grade for each assignment or assessment completed. Students earn points for a variety of activities, assessments, assignments, and behaviors that are accumulated throughout the grading period. Those points are then added together at the end of a term, and the teacher assigns a grade based on the total number of points earned (Marzano & Heflebower, 2011). With this traditional system, the receiver of the information has little to no true knowledge of what the student knows. The student could possess a limited understanding of all the content presented during the course, or understand some of the material very well but have significant learning gaps in particular areas. Consider for example a student who receives a final grade of 75% in a class that consisted of four exams. The possible distribution of grades was 90%, 30%, 90%, 90%, indicating a fairly strong student who did well for the most part but apparently had one area of great difficulty. However, the distribution could have been 67%, 67%, 100%, 67%. This distribution of scores presents a student who has struggled for most of the course, but had one shining unit in which he scored remarkably well. The final course average of 75% does not provide any insight into the actual performance of the student.

Issues of traditional grading methods. In this era of standards-based education, the problems with traditional grading practices are becoming more and more prevalent. One major problem associated with these grades is that different groups desire different approaches to calculating the grade, which leads to different interpretations of the grade. In a study of secondary teachers grading practices, Biberman-Shalev, Sabbagh, Resh, and Kramarski (2011) found a distinct difference in grading practices between teachers of

different subjects. When looking at 165 Israeli public high schools, the researchers found that teachers of mathematics preferred a performance-output style for grading, while science and language teachers preferred an input-effort style. Although there are clear differences between these subjects, a student receiving an A on a report card should indicate a mastery of the subject.

Secondary school teachers tend to give high grades to students who work hard in class, including effort as part of the grade (Brookhart, 1994). Over time this trend has resulted in grade inflation. Zirkel (2007) reported that students taking the SAT in 2000 scored no higher than their counterparts who took it in 1990, yet they earned average grade-point averages three tenths of a point higher in their high school classes. During the same period the number of students reporting an A average during high school had risen by 5%, but their SAT verbal scores dropped by 5 points and their SAT math scores dropped by 1 point.

A similar study was conducted on ACT scores between the academic years of 1989-90 and 1993-94 to confirm or rebuke the myth of grade inflation. By comparing the grades students earned in high school classes against their ACT composite scores, researchers discovered two trends. First, 75% of the students earning grade-point averages of 3.5 to 4.0 who attended a school in the top tenth of all high schools (as rated by ACT) made a composite score below 22, indicating a clear case for the existence of grade inflation. Second, when the 3.5-4.0 group of students at the top schools was compared to the same GPA band of students in the lowest school, the ACT scores differed by as much as seven points (Ziomek & Svec, 1995). Clearly, grading is not a uniform, consistent practice among teachers.

Grade inflation affects students' perceptions of their scholastic abilities; many enroll in college unprepared for the demands of collegiate level coursework. Jensen and Moore (2008) conducted a study of biology students at the University of Minnesota to discern how well their high school biology grades aligned with their college biology grades. On the first day of the college class, students were asked how well they felt high school had prepared them for college and what grade they expected to earn in the course. Results showed that none of the students earned lower than a C in high school biology, and collectively their high school GPA was an average of 3.3. Many of the students in the study indicated they never received a letter grade lower than a B in any course during their high school years, and 95% expected to make an A or B in the college course. At the conclusion of the college biology class, the researchers found the grade distribution for these students to be as follows: A=13%, B=27%, C=31%, D or F=29%. Their overall first year college GPA was 2.6. Grade inflation falsely indicates student achievement and leaves them ill-prepared for the reality of college.

Perceptions about grading. To address the issues created by the act of grading, one must consider the reason underlying the assigned grade. Brookhart (1993) conducted a study of classroom teachers to determine the meaning and values teachers associated with grades. In the study, 84 teachers were presented scenarios involving grading and asked for their responses in each situation. The scenarios involved questions of effort and ability, missing work, and student improvement. The analysis of the results revealed teachers' beliefs that grades are something students earn, much like a paycheck or compensation. Grades, at least from the perspective of the teachers, could be likened to currency, with students who worked hard receiving a better paycheck at the end of the

term. In this sense, the use of grades as a strategy for classroom management is apparent. Of interest in Brookhart's (1993) study is the evidence of a double standard. In one particular question, a student who performed better than average received a grade in line with her achievement, while a student who performed well below average often was rewarded with a better score if justification could be found. The perception of the teacher played a significant role in the final grade; students perceived as hard workers who did not cause discipline issues in class were awarded passing grades, even though they did not exhibit signs of academic achievement, while others who might have performed well on assessments and had a firm grasp of the content were punished with a lower score because of a perceived lack of effort.

Students value grades differently than teachers. Using focus groups at two undergraduate universities, Sanders and Anderson (2010) asked students about their perceptions of grades. The researchers learned that students believed grades did not adequately communicate their perceived level of success. Students were frustrated when they received Cs when they believed they had learned the content of the course; many indicated a belief that the only grade that signaled success was an A.

The differences between student and teacher perceptions of grades were studied further by Goulden and Griffin (1995). They asked teachers and undergraduate students to respond to two prompts, "What do grades mean to you?" and "Grades are like _____" (p. 110) in order to investigate the perception held by each group of subjects relative to grades. The results fell into eight basic categories; of particular interest was that 75% of the teachers had the same response. In their opinion grades served as a method of feedback to the students. Only 52% of the student responses indicated the same

perception. For 23% of the students, grades were emotional triggers, while 10% believed grades were a method of gate keeping. Different interpretations of the meaning of a letter grade from the stakeholders illustrate a significant problem with traditional grading methods.

While grade inflation appears to be rampant at one end of the spectrum, another problem emerges on the opposite end. Nationwide, 7,000 students drop out of school each day (Alliance for Excellent Education, 2010), in many cases because of failing grades (Bowers, 2011). Failing even one term of a math or English class could jeopardize a student's likelihood of graduating with his peers. Traditional grading methods leave students little room for error. Students are given assignments and assessments, with one attempt to score the best grade they possibly can. Once the teacher assigns a grade, it is final. Students who do not turn in an assignment typically receive a 0, which is averaged with all the other grades for the final mark. Students are not encouraged, and even may be forbidden, from working to improve their grades through reassessment (O'Connor, 2009).

Standards-based Grading

The shift in focus on standards in the classroom and the expectation of college for all have led many educators to reconsider the way grades are assigned. If all students are to meet a prescribed set of standards, and schools are held accountable for student performance, logically it would follow that the method whereby grades are given should be transformed to align with the standards. Prior to this age of accountability, students had more freedom to select courses of study that held their interest or were suited for their planned career choices. Students who were not exceptionally studious or who did

not like school could opt to take low level courses or many electives. Now all students, regardless of their aspirations for the future, are required to take a minimum of three years of math (typically two years of algebra and a year of geometry), four years of English, three of science, and three of social studies. In many districts, two years of a foreign language also is required (Education Commission of the States database, 2007).

To help all students reach these lofty and rigorous course requirements, a system based on scaffolding and support is crucial. Standards-based grading removes the connotation of finality that is present in traditional grading and replaces it with a feeling of hope for students. It is based on the idea that the purpose of assessment and a grade is to report accurately the achievement level of the student. In SBG, grades are based solely on demonstrating proficiency in content; elements of non-academic value (such as work ethic, behavior, and effort) are not included in the final grade report (Marzano & Heflebower, 2011; Stiggins, 2005).

In a truly standards-based classroom, a student no longer receives a single letter grade or percentage for a class. Instead of a grade for "math," the report card would contain individual marks for concepts within math. For example, in a first-year algebra class a student might receive individually reported grades in linear equations, systems of equations, quadratic equations, and polynomials, reassessing until mastery in each topic is reached. The student is encouraged to continue learning and practicing until it is possible to pass each concept with proficiency. Reporting grades in this manner shares explicitly the areas of strength and weakness of each student and removes from the calculation all non-academic factors (Guskey et al., 2011).

SBG shifts the focus of grades and education to real learning and student achievement. It is a first step in ensuring alignment in curriculum and testing standards, and an intense effort to remove "point-grubbing" (Cox, 2011) and "grade fog" (Deddeh, Main, & Fulkerson, 2010) from student grades. SBG provides a medium to assist in eliminating grade inflation and grading inequities, as well as offering students a sense of hope and a structure designed to foster success.

Current SBG initiatives. SBG has been relatively well-received in the elementary grades, with entire schools adopting the practice completely, including revising the report card format parents receive each grading period. The North Spencer School Corporation in Indiana (Tassell, Kemp, Litkenhus, & Schriefer, 2006) adopted SBG in its elementary schools (grades K-6) after reviewing student performance on the state-mandated Indiana Statewide Testing for Educational Progress Plus (ISTEP+) standardized test and student grades, specifically focusing on the students who did not pass the test. Of the 101 students who failed at least one section of the ISTEP+ test in the fall 2001, 85 earned an A or B in math or reading. This discrepancy between classroom grades and external assessment scores motivated the district to not only seek close alignment between the content standards and classroom instruction, but also to seek a reporting method that would better inform parents of their child's true achievement level. The results of implementing SBG in this district were very successful. Parents felt better informed about their child's performance, thus improving the lines of communication between the school and the home. Additionally, SBG appeared to address the issue of grade inflation. In 2001-2002, 53% of students who earned an A or B failed the English/language arts portion of the ISTEP+. After implementing SBG in 2004-2005,

only 32% of students who were reported either at or above grade level failed the same portion of the test. In math, the percentage of failures for A/B students was 32 in 2001-2002; that percentage decreased to 17% in 2004-05 (Tassell et al., 2006).

A similar study confirmed the same findings in elementary schools in Northern Georgia. Hardegree (2012) analyzed fifth-grade students in eight elementary schools in North Georgia to determine whether grades on SBG report cards would provide an accurate measure of the grades received on a standardized criterion-referenced test. Results of this study found that a comparison of students who met, did not meet, or exceeded classroom standards had significantly different means on the state exam in mathematics and reading. Students who received classroom grades of "meeting standards" scored higher on the test than those who received grades of "not meeting," while students with grades of "exceeding standards" outscored both of the other groups. Hardegree's study suggested SBG marks are an appropriate predictor of performance on external assessments.

Not all studies of SBG show improvement in student performance. In contrast to the studies by Tassell et al. (2006) and Hardegree (2012), Craig (2012) analyzed student achievement scores on the mathematics portion of the fourth-grade Massachusetts Comprehensive Assessment System. Comparisons of SBG to non-SBG schools suggest no difference in the growth or the performance levels of students, regardless of the school type or report type. However, the study found that components of SBG may have a positive academic impact on at-risk students.

Studies of high school students in SBG classrooms appear to be sparse. In a quasi-experimental study, Hartnell (2011) investigated the achievement of students in

American government classes, comparing traditional teaching methods to methods using a standards-based curriculum, differentiated instruction, and performance-based assessments. Hartnell found that, while the students who were exposed to SBG performed better than their non-SBG counterparts, the difference was not statistically significant. Hartnell's study has serious limitations that should be considered when gauging the accuracy of his results. Most prominently is the fact that the two groups of students were taught by two different teachers. The lack of a significant difference could be due to teacher personality and instructional style rather than the experimental variable. In addition, standards-based assessment and differentiated instruction are not the same and attempting to analyze both of them in the same study could have confounding effects. Nonetheless, the study illustrates the need for further testing and research.

Effects of Grading

The effects of grading on students are widespread. Students who generally perform well in school tend to view assessment and grading as evidence of their success. They are likely to seek challenges, take risks, and interpret assessment as opportunities to gain feedback. They tend to persevere during setbacks and accept responsibility for their results. Each of these feelings leads to more success and a positive cycle develops. In contrast, students who do poorly in school view assessments as evidence of their failures. Instead of accepting new challenges and taking risks, these students feel hopeless and seek the easiest options. Rather than developing perseverance, they learn to retreat, avoiding initiative. These students fall into a negative cycle leaving them in a state of frustration, fear, and defeat (Stiggins, 2007).

Grading and its effects on students is a central component in the conversation about standards-based education. Ultimately, students are the real users of assessment information (O'Connor, 2011), and educators need to consider how grading impacts the students. Since students spend a considerable portion of their school career involved in assessment of some type, it behooves educators to understand how this manifests itself in student beliefs and behavior. Many of the issues present in traditional grading methods have negative effects on students and can be corrected by the effective implementation of SBG (O'Connor, 2011).

Student Motivation

Contrary to many opinions, traditional grading practices have adverse consequences on student motivation. There are two types of motivation — intrinsic and extrinsic. Intrinsic motivation is the act of doing something for the inherent enjoyment or interest, while extrinsic motivation is the act of doing something because of the resulting outcome (Ryan & Deci, 2000). Assigning grades to student work as a motivational factor can be a problem for the students. Crooks (1988) reported an analysis of three studies that show that, when a grade is assigned to an activity that previously was voluntary, student interest declined and they were less likely to return to the activity. If students were working on a project or activity because of interest (intrinsic motivation), they were likely to persevere, even when the task became difficult. In contrast, students who were working on a project because they were going to be rewarded chose only the easiest tasks. In addition, the extrinsic student group was driven by answers and often looked for short cuts, whereas the intrinsic group sought deeper knowledge and more meaningful approaches to the projects.

Extrinsically motivating tasks can lead to the development and enhancement of intrinsic motivation. The task for teachers is to ascertain out how to make grading and evaluation purposeful so that grades become a medium for improving intrinsic motivation. If students can develop a sense of belonging and connectedness to a task they would not normally complete, the likelihood of completion increases. This desire to please another person serves as an extrinsic motivation factor and is the first step in turning extrinsic into intrinsic motivation (Ryan & Deci, 2000).

Second, if students are to turn the grading process into an intrinsic motivator, they need to feel a sense of competence toward the task they are being asked to do (Ryan & Deci, 2000). Through the use of learning targets and instruction that is aligned to the assessment, students should develop the feelings of competence necessary to succeed in the classroom. Students who believe the goal of an assessment is to provide helpful and informative feedback, rather to control their behavior, tend to develop a deeper sense of intrinsic motivation for the course (Crooks, 1988).

Student Self-efficacy

Students must deal with the discrepancies of grading methods as often, if not more, than teachers. Grades should be indicators of academic success, but often do not reveal what a student might know. Students at the undergraduate level of education must work to make sense of the grades they receive, often trying to find personal growth and instances of genuine learning in spite of low grades (Sanders & Anderson, 2010).

Grades can play a key role in affecting a student's self-efficacy. A study by Shim and Ryan (2005) investigated how grades in the classroom affected students' self-perceptions and their intrinsic values. By surveying 361 college students using a 7-point

Likert-type scale, they found that higher grades correlated with higher self-efficacy; in addition, students who received higher grades generally felt more intrinsic value from the class than those who received lower grades. The results of this study also suggest that, at least for college level students, mastery goals (those that focus on understanding and competence) support the development of self-efficacy. In light of the current trends in education, this study provides some support for the use of learning targets and standards-based grading in the classroom.

Student self-efficacy may be fostered best through the use of long- and short-term goals in conjunction with mastery learning procedures. Students need opportunities to improve their self-perception with real achievement results (Crooks, 1988). Self-monitoring of growth and development on learning targets encourages students to set personal goals, while mastery of targets provides students with authentic success (Moss & Brookhart, 2012).

Student Performance

In a meta-analysis of 64 research studies conducted between 1968 and 1990, Anderson and Pavan (1993) reviewed comparisons of student achievement in non-graded schools compared to their counterparts in similar graded schools. When looking at the studies that used standardized achievement tests to make comparisons, only 9 (out of 94) resulted in graded schools outperforming the non-graded schools. The dates of the research used in this analysis coincide with the time in educational history when traditional letter grading based on percentage was the norm. It is fairly safe to assume the schools that used grading methods were most likely grading with a traditional method,

and the results of this analysis suggest that perhaps traditional grading does not have a positive effect on student performance.

To further review the relationship between classroom grades and achievement on high-stakes tests, Haptonstall (2010) conducted a study to determine the correlation between classroom scores and end-of-course (EOC) assessment marks of students in grades 6 through 12 in a sample of Colorado school districts. The focus of this study was to see how closely classroom grading policies mirrored the achievement of students on the state assessment and to determine if the grading practices implemented in classrooms adequately measured the students' level of achievement. Although the study was not specifically designed to compare SBG to traditional grading, one school district in Colorado implemented SBG unilaterally and allowed for a comparison of SBG to traditional grades. Haptonstall compared this district to non-SBG districts and found a significantly higher correlation between classroom grades and end-of-course assessments for those students in the SBG environment. In addition, the EOC exam grades were higher for all populations of students who attended the SBG district. Not only did this study find SBG students performed better, but also it revealed numerous grading discrepancies in non-SBG schools for ELL students, low SES students, and Hispanic students. Pupils in these populations often scored higher in the classroom than on the EOC exam, possibly giving them a false sense of accomplishment prior to the EOC exam.

Summary

This chapter has established the theoretical, historical, and contextual basis for the use of standards-based grading in the classroom. A review of the literature reveals that

education is taking a new approach to learning and assessment, driven by accountability. Educators need new methods and practices for grading students, methods that encourage, support, improve, and enhance the learning process. SBG provides one possible alternative to teachers that can assist in aligning classroom marks to external assessments. SBG certainly has changed the way teachers report grades on report cards (Tierney, Simon, & Charland, 2011), with teachers indicating improved student performance based on classroom grades (Guskey et al., 2011) when SBG is implemented. While improved performance on teacher-made exams is a start to improving achievement levels, school systems will not make advances toward meeting the requirements of educational reform mandates from the government if students cannot perform better on externally assessed instruments. This study will review the effects of SBG on student performance on end-of-course assessments to determine the potential of SBG to improve student achievement. The next chapter will outline the methodology used and provide the contextual setting for the study.

CHAPTER III: METHODOLOGY

Standards-based grading (SBG) is gaining momentum in the school system as a result of the shift toward increased testing and accountability. The emphasis on preparing all students for post-secondary education has caused educators to rethink the purposes of grades as well as the methods by which those grades are determined. Many teachers have begun using it with the hope of improving student achievement by emphasizing mastery learning as opposed to point-collecting. The basic premise of SBG is that grades should reflect what a student has learned rather than serving as a tally to the amount of points a student has accumulated during a period of time (Brookhart, 2011). It follows that a shift in student focus, from point accumulation to mastery learning, should result in improved measures on student achievement. While some evidence supports the use of SBG to improve student achievement in elementary school (Tassell et al., 2006; Hardegree, 2012), the research for its use at the high school level is sparse. Given the importance districts and political leaders are placing on student performance, it is essential to consider reform efforts on grading practices from an empirical perspective. This study seeks to determine any association between students who were graded with an SBG approach and their scores on posttests or on end-of-course (EOC) examinations.

Research Questions

The focus of this quantitative research study was to determine whether a difference exists in student achievement on state mandated end-of-course (EOC) exams and on classroom administered posttest results between students who were members of a classroom that implemented standards-based grading (SBG) and those who were members of a traditionally-graded class. The study was guided by the central research

question: Does standards-based grading affect student performance in Algebra 2? Under this central research question, the specific research questions addressed in this study include the following:

1. How do the grading method and the student achievement level impact the Algebra 2 end-of-course assessment?
    a. How does the grading method (standards-based grading and traditional grading) impact the results of the Algebra 2 end-of-course assessment?
    b. How does the student achievement level (honors and regular) impact the results of the Algebra 2 end-of-course assessment?
    c. Is there interaction in student performance on the Algebra 2 end-of-course assessment between the grading method and the student achievement level?
2. How do the grading method and the student achievement level impact the classroom posttest results?
    a. How does the grading method (standards-based grading and traditional grading) impact the classroom posttest results?
    b. How does the student achievement level (honors and regular) impact the classroom posttest results?
    c. Is there an interaction in student performance on the classroom posttest results between the grading method and the student achievement level?

Research Design

The design of this quantitative study was based on a causal-comparative design, or ex post facto research. In this design, the researcher looks for an association between

variables after an event has taken place. It is similar to correlational research, in that the researcher seeks to discover a relationship between independent and dependent variables; however, a causal-comparative design does not assign subjects to treatment or control groups. Instead, subjects are grouped based on a defined characteristic (or treatment), and statistical testing is implemented to determine whether a difference exists between groups who experienced the treatment and groups who did not. A design of this nature was warranted because it was not possible to randomly assign students to each of the grading strategies (Brewer & Kuhn, 2010).

Participants

The administrative leader of the math department had approached one math teacher approximately nine months before the start of this analysis to propose a study of standards-based grading. The administrator had observed its implementation in middle schools in the district and, after attending a professional development seminar regarding SBG, was interested in its potential for the school. The math grades of the high school students at that time were problematic, with a low percentage of students able to achieve a state-identified benchmark score on EXPLORE, PLAN, and ACT exams. The belief was that SBG could be a mechanism for improving student achievement. As a result of this conversation, two teachers agreed to conduct this pilot study to discover whether any improvement in EOC exam scores could be obtained using SBG.

Two teachers and four classes of second-term Algebra 2 students participated in this study. The teachers each volunteered to implement SBG in one class and maintain a traditional grading method in another. Teacher One was in the 17th year of teaching and held a master's degree in education with a minor in mathematics; Teacher Two was in the

16th year of teaching with a master's degree in mathematics and a minor in education. The instructors had taught together for 14 years and had taught Algebra 2 for 17 years and 13 years, respectively. Both teachers have held National Board Certification since 2002. Prior to agreeing to participate in this study, both teachers had attended a summer workshop on the topic of SBG, specifically regarding writing effective learning targets and aligning assessment to those targets. Each had implemented the practice of posting learning targets in all classes and had started organizing assessments by learning targets, grouping questions and providing indicators for students so they could identify which learning targets were associated with each group of questions. After agreeing to assist with the study, both teachers read updated literature regarding SBG and began to detail a plan of implementation to ensure that the core ideals of SBG were maintained. In the interest of full disclosure, Teacher Two was the researcher and author of this dissertation and study.

A convenience sample of high school students enrolled in the second half of Algebra 2 were selected as the participants. The public school is located in a Midwest city with a population of 57,600 and serves approximately 1,300 students. The student body is 87.1% Caucasian, with just over 50% reporting free or reduced lunch status (Kentucky Department of Education, 2013). A total of 107 students participated in the study, 46% male and 54% female. Approximately 40% of the students qualified for free or reduced lunch status, and 17% reported a non-Caucasian ethnicity.

To ensure that no statistically significant difference existed between the SBG classes and the traditionally-graded classes prior to the start of the pilot, an independent sample t-test was conducted using a teacher-constructed multiple-choice pretest. Students

in the traditionally-graded classes scored only slightly higher (M = 11.08, SE = .406) than the SBG students (M = 10.94, SE = .479). The difference of these pretest score means was not significant $t(100) = -.219$, $p > .05$; neither group had an advantage at the start of this investigation.

Measures

The existing data set of EOC exam scores were used as the measure of student performance on the EOC test. The exam was divided into two 38-question sections; students were allowed 45 minutes for each. For the computer-based exam, students were provided a TI-84 graphing calculator and extra paper. The guidance office provided exam schedules, and students took the test during a normal school day in a classroom proctored by a teacher or guidance counselor. Scores were calculated by the testing vendor, and students were given their score 24 hours after completion of the exam. This score was weighted as 10% of the students' final grade. The scale score range for EOC exams was 125-175. Of the sample, five students were deemed ineligible by the school to sit for the EOC exam due to failing at least one part of the Algebra 2 course. Those students have not been reported in any statistical analysis of the EOC results.

To determine student performance at the beginning and end of the term, a test was created using a database of questions published by Quality Core, the manufacturer of the EOC exam. The questions were selected from the test bank published by Quality Core using a stratified random sampling method to ensure that all topics and all depth of knowledge levels were represented. Seven questions from each topic (Table 1) were selected, generating a 35-question test. Students were given 60 minutes to complete the test and were provided a TI-84 graphing calculator. The pretest and posttest were

administered during the students' Algebra 2 class and were paper based. Scores for these assessments were reported as the percentage of correct responses.

Table 1
Distribution of Topics for Posttest

Category (7 questions from each)
Linear Functions
Number sense, quadratic functions, and matrices
Polynomial functions
Non-polynomial functions
Probability, sequences, and series

(Quality Core, 2012, p. 1)

Due to late schedule changes and illness, five students did not take the pretest. Their scores were not used in the t-test to ensure equitable starting populations; however, their posttest scores were used, as they had been enrolled in the course for more than 10 of the 12 weeks. Two students did not take the posttest due to extended absences. Those scores were omitted from the statistical analysis of that measure but were included in the analysis of the EOC exam.

Reliability and Validity

Two exams were used as dependent variables for this study, the EOC test and a posttest. The EOC exam was authored by the manufacturers of the ACT and has undergone extensive testing for validity and reliability (Quality Core, 2010). The posttest used for this study was created using a set of published benchmark assessment questions issued by ACT and the writers of the EOC exam to ensure its validity and reliability.

To maintain the integrity and usefulness of this study, every effort has been made to maintain the highest level of internal validity. However, potential threats to the internal validity were noted that cannot be removed and must be taken into account when interpreting the results. The actual act of testing is a threat to internal validity, with performances affected by expectations of the researcher or teacher, practice, and memory of earlier responses (Pedhazur & Schmelkin, 1991). Since the EOC exam was only administered at the end of the course, practice and memory would not have been potential threats; however, the expectation of the teacher may have affected some students. Since the students were required to participate in a pretest and a posttest, expectations, memory, and practice could have been factors of the scores obtained on the posttest.

The study was conducted during a 12-week course, thereby minimizing the maturation of students as well as the time between the pretest and posttest. Although experiences outside the classroom were most likely different for each student, the assumption can be made that the in-class experiences for the students were very similar. Teacher One taught both sections of regular Algebra 2, and Teacher Two taught both sections of the honors level; each was diligent about offering the same learning activities and teaching methods to their respective classes.

Regardless of the intentionality on the part of the teachers to ensure equitable experiences to all of their students, one threat to internal validity became apparent after only a few weeks into the course. Students in the traditionally-graded courses had the perception that they were being cheated or dealt with unfairly when they learned that their counterparts in the SBG classes were allowed and encouraged to retake assessments

to improve their scores. In some cases, the emotional disconnect created by this difference may have played a role in the student's overall effort and assessment results.

Procedures

The school operated on a five-period day on a trimester schedule, and most students were new to each instructor at the beginning of the course. A small proportion (n=12) had participated in the first half of Algebra 2 or in a probability and statistics class with the same teacher to whom they were assigned for the study. Classes were approximately 70 minutes long on Mondays and Fridays and approximately 60 minutes long the remaining days of the week. The school was a one-to-one laptop school (each student was assigned a personal laptop computer for use at home and at school); in addition, both classrooms were equipped with classroom sets of graphing calculators (TI-84's), an interactive calculator system (TI-Navigator), and Smartboards.

Both teachers randomly selected one class to be assessed using SBG and one to be assessed with traditional grading methods; sample sizes of each section are reported in Table 2. Teacher One taught only regular classes during the term, and Teacher Two taught only honors classes. The household of all students received an informative letter (Appendices C and D) detailing the grading method that would be used and soliciting any questions or comments from parents. Because SBG was such a radical departure from which the students and parents were accustomed, a letter of explanation was deemed necessary. To prevent parental involvement from becoming a confounding variable, the parents of students in the traditionally-graded classes also were sent a letter of welcome and introduction to the course.

Table 2
Assignment of Students to Teacher

Teacher	Grading type	Level	Sample size
One	Traditional	Regular	28
One	Standards-based	Regular	26
Two	Traditional	Honors	26
Two	Standards-based	Honors	25

Within the first three days of each class, the students were administered the pretest, which was scored using a Scantron machine. Teacher One entered the data into a spreadsheet, recording the level of the student, the class they attended (to allow a grading method to be tracked), an exam score, and other demographic information. This information was placed in a secure location until the end of the term and students were informed of their results.

For the duration of the term (approximately 12 weeks), the instructors taught both sections of their classes using the same instruction methods, activities, and resources in each of their respective classes. To ensure equitable activities, the instructors met weekly for the purposes of planning and sharing resources. For students in the traditional courses, quizzes were given periodically during each unit, and a unit test was given at the conclusion of the unit. Per the directive of the district, learning targets for traditional classes were posted daily. Students were not required to re-test but in certain circumstances were given the opportunity to do so, namely, when a student average was below a passing grade. If these students were permitted to reassess, they were required to re-take the entire quiz or test they were attempting to improve.

For students in the SBG classes, learning targets (LT) were provided, and students were instructed on a simple method for tracking their own performance. At the end of each week, SBG students took a learning target quiz. Each was created by the instructors and covered three to five learning targets, with a minimum of three questions per target. The questions ranged in difficulty level, and the instructors scored each LT using the grading rubric presented in Table 3. After the weekly quiz was graded and returned, students marked their progress on their student tracking form. To promote long-term retention of skills, students were required to assess each LT twice (typically on consecutive weeks), and both scores were recorded and entered into the grade book. Students who earned a grade of less than C on any LT were required to remediate and retest. The retest grade replaced the lowest grade in the grade book. To obtain a final average, all LT scores were averaged and weighted at 90% of the students' final course grade; the EOC exam score constituted the remaining 10% of the course grade.

Table 3
Scoring Rubric for Learning Targets (LT)

Letter grade	Numeric grade (for reporting purposes)	
A+	10	Two examples of a completely correct solution to the problem, with no errors, that demonstrates correct use of concepts and skills.
A	9	A completely correct solution to the problem, with no errors, that demonstrates correct use of concepts and skills.
B	8	A solution to the problem that demonstrates the concepts and skills are known and can be used correctly, but non-fatal errors are present.
C	7	A solution to the problem that demonstrates the concepts and/or skills are known, but fatal errors prevent the correct solution from being found.
Not Yet Met	5	A solution to the problem that demonstrates the concepts and/or skills are not known to a degree of usefulness.

Table 4 summarizes the differences between the SBG classes and the traditional classes.

Table 4
Summary of SBG Classes vs. Traditional Classes

Standards-based Classroom	Traditional Classroom
12 weeks of class	12 weeks of class
Learning Targets posted daily	Learning Targets posted daily
2 required, unique assessments per LT, administered approximately 5 days apart	1 required assessment per LT
3rd assessment option to demonstrate mastery	
Assessed frequently — at least once every 2 weeks	Assessed at the end of unit — 4 assessment opportunities total
"Lots of grades" (student opinion)	"Not very many grades" (student opinion)
Grades = A, B, C, NY	Traditional percentage grades
Students tracked grades	No grade tracking

Data Analysis

Research approval from the Institutional Review Board (IRB) of Western Kentucky University was requested and obtained in January 2013 prior to all data collection. Administrative approval from the school district was secured as well. IRB approval and district support letters are included in Appendices A and B.

Upon completion of the course and all assessments, the researcher was provided with a Microsoft Excel spreadsheet containing the following information for each

student: gender, ethnicity, socio-economic status, pretest score, posttest score, EOC exam score, grading method (SBG or traditional), and level (honors or regular). The researcher imported the data into IBM SPSS Statistics software for the purpose of analysis.

Since the independent variables are categorical (assessment type and achievement level) and the dependent variables (test scores) are measured on an interval ratio, two-way ANOVA testing was used to determine any significant differences in test scores between SBG classes and traditionally-graded classes. In addition, two-way ANOVA testing was used to determine any interaction between the independent variables that could indicate a unique effect resulting from a combination of the independent variables (Gravetter & Wallnau, 2009). Provided the sample sizes are roughly equal, ANOVA testing can be conducted regardless of the normality of the distribution (Field, 2009). Each class of students had a sample size of approximately 25, allowing ANOVA testing to serve as an appropriate statistical tool.

To answer the first specific research question: How do the grading method and the student achievement level relate to the Algebra 2 end-of-course assessment? ANOVA testing was used to determine whether a difference existed between the EOC exam performance of students who received traditional grading and those who received SBG. To address the second research question: How do the grading method and the student achievement level relate to the classroom posttest results? ANOVA testing was conducted to determine whether a significant difference was found between the mean achievement scores on the post assessment of the students who received standards-based grading and those who received traditional grading methods.

To determine outcomes that may be considered statistically significant, a confidence level of $\alpha = .05$ was used. This confidence level separates the most unlikely 5% of the outcomes from the remaining 95%. Any results obtained with a *p*-value of less than .05 can be considered significant and should be considered for further research or testing (Gravetter & Wallnau, 2009).

Summary

This chapter has served to outline the research design of the study, detailing the demographics of the participants, providing a setting for the nature of the pilot study, offering justification for the statistical analysis that was conducted, and discussion potential limitations to the application of the results. Chapter IV will present the results of the study and the statistical tests; Chapter V will discuss those results in detail and share the potential implications of the findings.

CHAPTER IV: RESULTS

The culture surrounding grading structures in American education is changing and standards-based grading (SBG) is beginning to sweep the nation. Teachers in all grade levels and content areas are adapting the framework of SBG as a means of transforming the way students earn grades and the way those grades are communicated to parents. SBG offers all students the opportunity to demonstrate mastery of learning objectives, or targets, through reassessment; additionally, SBG can alter how teachers, students, and parents interpret student achievement. While SBG appears to be a more equitable grading practice, the impact of its implementation on high stakes testing remains largely unknown. This study seeks to answer the question of the effectiveness of SBG to improve student performance on summative assessments in Algebra 2.

Method of Analysis

The independent variables of this study were achievement level (honors or regular) and grading method (traditional or SBG); the dependent variables were end-of-course (EOC) test scores and classroom posttest scores. Two separate runs of two-way ANOVA testing allowed for a comparison of the means of exam scores, as well as any interactions that could have resulted from a combination of the independent variables. Figure 3 illustrates the potential associations between the variables.

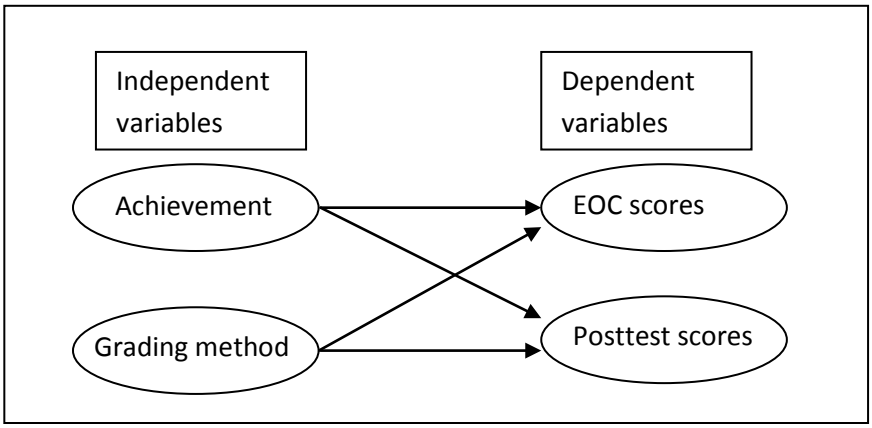

Figure 3. Variable assignments and potential associations between grading method and test scores

Research Questions

The central research question that guided this study is: Does standards-based grading affect student performance in Algebra 2? Under this central research question, the specific research questions addressed in this study include the following:

1. How do the grading method and the student achievement level relate to the Algebra 2 end-of-course assessment?
    a. How does the grading method (standards-based grading and traditional grading) impact the results of the Algebra 2 end-of-course assessment?
    b. How does the student achievement level (honors and regular) impact the results of the Algebra 2 end-of-course assessment?
    c. Is there interaction in student performance on the Algebra 2 end-of-course assessment between the grading method and the student achievement level?
2. How do the grading method and the student achievement level impact the classroom posttest results?

a. How does the grading method (standards-based grading and traditional grading) impact the classroom posttest results?
b. How does the student achievement level (honors and regular) impact the classroom posttest results?
c. Is there an interaction in student performance on the classroom posttest results between the grading method and the student achievement level?

Results for Research Question 1

To address the first research question: Does standards-based grading affect student performance in Algebra 2?, the end-of-course test results for four Algebra 2 classes were analyzed. The classes were divided into achievement levels (honors or regular) and grading methods (SBG or traditional). The sample sizes of the four classes are presented in Table 5.

Table 5
Distribution of Students for Grading Method and Achievement Level

	Traditional	SBG
Honors	27	25
Regular	25	25

Summary statistics for the EOC test scores are presented in Table 6. The overall average score for all students on the EOC test was 147.35, with honors students averaging just slightly higher (148.10) and regular students averaging a bit lower (146.84). Students could score between 125 and 175.

Table 6
Descriptive Statistics of the EOC Assessment Scores

Achievement Level	Grading Method	*Mean*	*SD*	*min*	*max*	*N*
Honors	SBG	148.12	3.72	139	156	25
	Traditional	148.07	3.23	140	155	27
	Total	148.10	3.44	139	156	52
Regular	SBG	147.12	2.93	141	153	25
	Traditional	146.56	2.80	139	152	25
	Total	146.84	2.85	139	153	50
Total	SBG	147.62	3.36	139	156	50
	Traditional	147.35	3.10	139	155	52
	Total	147.48	3.21	139	156	102

Two-way ANOVA testing requires independent observations, approximately equal variances, and a quantitative measurement scale for the dependent variable. For this study, student test results were independent of each other, and both EOC scores and posttest scores were based on an interval scale. Based on the Levene's test, the homogeneity of the variance assumption was met ($F = .699$, $p = .555$).

Research question 1a asks: How does the grading method impact the results of the Algebra 2 end-of-course assessment? The grading method was not found to be a significant factor on EOC scores ($F(1,98) = .229$, $p = .633$). Based on these results, neither grading method was found to impact students' performance on the EOC test. The mean score of both groups was approximately the same.

To address the first research question, a two-way ANOVA test was conducted using EOC scores as the dependent variable. The results are presented in Table 7.

Table 7
The Results of the Two-way ANOVA for the EOC Assessment

Source	*SS*	*df*	*MS*	*F*	*p*
Achievement Level	40.249	1	40.249	3.947	.050
Grading Method	2.338	1	2.338	.229	.633
Achievement Level × Grading Method	1.683	1	1.683	.165	.685
Error	999.292	98	10.197		
Total	1043.461	101			

R Squared = .042 (Adjusted R Squared = .013)

Research question 1b asks: How does the student achievement level (honors and regular) impact the results of the Algebra 2 end-of-course test? The results indicate a statistically significant difference on EOC test scores between honors and regular students at the $\alpha = .05$ level ($F(1,98) = 3.947$, $p < .05$). Based on these results, honors students perform significantly better than regular students in the EOC assessment.

Research question 1c asks: Is there an interaction in student performance on the Algebra 2 end-of-course assessment between the grading method and the student achievement level? Analysis of the results of the ANOVA testing indicate no significant interaction between achievement level and grading method ($F(1,98) = .165$, $p = .685$). SBG was not found to benefit one achievement level any more than the other.

Results for Research Question 2

To explore the second research question: How do the grading method and the student achievement level impact the classroom posttest results?, a two-way ANOVA test was conducted using achievement method and grading method as independent variables and posttest scores as the dependent variable. A sample size of $n = 105$ was obtained using four different Algebra 2 classes. Scores on the posttest ranged from 4 to 20 (out of a maximum possible of 35), and the average of all scores was 12.25 with a standard deviation of 3.35 (see Table 8).

Table 8
Descriptive Statistics of the Posttest Scores

Achievement Level	Grading Method	*Mean*	*SD*	*N*
Honors	SBG	12.56	2.987	25
	Traditional	14.65	3.199	26
	Total	13.63	3.243	51
Regular	SBG	10.96	3.194	26
	Traditional	10.93	2.680	28
	Total	10.94	2.910	54
Total	SBG	11.75	3.168	51
	Traditional	12.72	3.466	54
	Total	12.25	3.345	105

Based on the Levene't test of equality, the homogeneity of the variance assumption was met (F = .440, p = .725), thus allowing for ANOVA testing (See Table 9).

Table 9
The Results of the Two-way ANOVA for the Posttest

Source	*SS*	*df*	*MS*	*F*	*p*
Achievement level	185.683	1	185.683	20.410	.000
Grading method	27.826	1	27.826	3.059	.083
Achievement level x Grading Method	29.635	1	29.635	3.257	.074
Error	918.863	101	9.098		
Corrected Total	1163.562	104			

R Squared = .210 (Adjusted R Squared = .187)

Research question 2a asks: How does the grading method (standards-based grading and traditional grading) impact the classroom posttest results? Using a significance level of α = .05, grading method (F(1,101) = 3.059, p = .083) was not statistically significant on posttest results. As with the EOC test, neither grading method was more beneficial than the other on the posttest.

Research question 2b asks: How does the student achievement level (honors and regular) impact the classroom posttest results? Analysis of the ANOVA testing indicates that achievement level (honors or regular) was statistically significant (F(1,101) = 20.410, p <.05) for student performance on the posttest. The mean of the honors students was significantly higher than the mean of the regular level students.

Research question 2c asks: Is there an interaction in student performance on the classroom posttest results between the grading method and the student achievement level? ANOVA results indicate no significant interaction between the achievement level and the grading method ($F(1,101) = 3.257$, $p = .074$).

A post hoc ANOVA test was conducted using only the 16 questions on the posttest that tested learning targets directly taught during the last half of the Algebra 2 course when the standards based grading was implemented (see Table 10). Achievement level was a significant variable ($F(1,101) = 9.495$, $p = .003$); but grading method ($F(1,101) = .300$, $p = .585$) was not found to be significant. There also was no significant interaction between grading method and achievement level ($F(1,101) = 1.737$, $p = .190$).

Table 10
The Results of the Post HocTtwo-way ANOVA

Source	*SS*	*df*	*MS*	*F*	*p*
Achievement level	32.304	1	32.304	9.495	.003
Grading method	1.021	1	1.021	.300	.585
Achievement level * Grading method	5.910	1	5.910	1.737	.190
Error	343.611	101	3.402		
Corrected Total	383.429	104			

a. R Squared = .104 (Adjusted R Squared = .077)

Summary

This chapter has presented the results of the statistical analysis. The results of the two-way ANOVA analysis for the two outcome variables show that honors students perform better than regular students on the state-issued end-of-course examination and on the teacher-written posttest. However, no significant impact was found between grading

method and achievement level on EOC scores or posttest scores. Chapter V will consider the implications of these findings and discuss areas of potential further research.

CHAPTER V: CONCLUSIONS

This study has examined the potential effect of standards-based grading (SBG) on student performance in Algebra 2, specifically on the end-of-course (EOC) exam and a course posttest. Educational reform initiatives, such as No Child Left Behind and Race to the Top have placed teachers and education in a positive spotlight nationwide with individuals searching for methods and strategies that can be implemented to raise student achievement. SBG has surfaced as one reform idea with the potential to increase student performance by shifting the focus away from classroom grades. Under SBG initiatives, students can view grades as progress reports, providing students with an indication of their achievement level during a course and alerting them to the objectives and skills of the class that are in need of attention and improvement. SBG removes the connotation of finality that has become associated with a traditional classroom average (Cox, 2011).

SBG requires students to achieve a prescribed level of proficiency, but allows opportunities to continuously reassess until that level of proficiency has been demonstrated. The purpose of this study has been to test the effectiveness of SBG to raise student achievement by addressing a central research question: Does standards-based grading affect student performance in Algebra 2? Under this central research question, the specific research questions addressed in this study include the following:

1. How do the grading method and the student achievement level relate to the Algebra 2 end-of-course assessment?
    a. How does the grading method (standards-based grading and traditional grading) impact the results of the Algebra 2 end-of-course assessment?

b. How does the student achievement level (honors and regular) impact the results of the Algebra 2 end-of-course assessment?

c. Is there interaction in student performance on the Algebra 2 end-of-course assessment between the grading method and the student achievement level?

2. How do the grading method and the student achievement level impact the classroom posttest results?

    a. How does the grading method (standards-based grading and traditional grading) impact the classroom posttest results?

    b. How does the student achievement level (honors and regular) impact the classroom posttest results?

    c. Is there an interaction in student performance on the classroom posttest results between the grading method and the student achievement level?

Discussion of Findings

A comparison of both EOC and posttest scores showed that students enrolled in honors level courses outperformed their counterparts in regular classrooms. However, no significant difference existed between SBG students and traditionally-graded students at either level. These results were surprising to the researcher, given the current research and literature base that supports the implementation of SBG.

Connection to Theoretical Framework

The theory of formative assessment provided the primary framework for this study. Three components are necessary for successful implementation of formative assessment. Students must first know what is expected of them, have the ability to self-

assess their current ability level for those expectations, and have a plan of action whereby they can close the learning gap (Sadler, 1989).

While the researcher as teacher was offered as a possible limitation to the study, it served as a valuable opportunity for gaining insight into the surprising results. By addressing each of these components identified as necessary for a successful SBG classroom, much different results were expected at the conclusion of the study.

The first element, that students know what is expected of them, was accomplished through the use of learning targets, which can be found in Appendices E and F. Both SBG and traditionally-graded classes were given learning targets (LT) per district directions; therefore, it is difficult to ascertain the impact of this element of the SBG classroom. Students claim the LTs made the course content easier to navigate, but this was true of both traditional classes and SBG classes.

The second critical element of SBG is for students to have the ability and opportunity to assess their performance on each standard and determine what they need to do to improve (Sadler, 1989). For this study, students in the SBG classes were required to keep a record of each learning target and their subsequent scores on the in-class assessments. Students were provided time upon the return of each graded assessment for the purposes of recording and reflecting on how they could improve. Although this was stressed as important to the students, the feeling of the teachers involved was that the students did not take this exercise seriously, and most never seemed to grasp the full intent of this practice. If this is, in fact, a key element for a successful SBG experience, it would seem that a much more intentional, focused effort is necessary to help the students

genuinely reflect on their learning. For most students, individual grades did not motivate them to improve; rather, they were still motivated by the overall average score.

The final element of the theoretical framework requires a clear plan of action for students to decrease the gap in their knowledge (Sadler, 1989). This study was designed with an action plan that required students to spend a minimum of 30 minutes of independent time (outside of class) working on the learning target they wished to reassess. This could be time spent with the teacher, in after-school tutoring services, watching tutorial videos, or any number of activities the students felt would benefit them. To avoid cramming and short-term memorizations, students were not allowed to revise their learning target and reassess on the same day; all retests had to be completed at least one day after correcting, studying, and receiving help. After students were instructed in the grade tracking system and presented with the variety of options available to them to assist with closing gaps and helping to improve learning target cores, it became apparent that they needed more guidance when developing their plan of action. Most lacked the ability or motivation to effectively help themselves improve. Even when resources (such as practice problems and video examples) were made available to them, they seemed to resist completing the work. They viewed it as punishment rather than seeing it for what it was meant to be, the opportunity to learn.

Discussion of Research Question 1

Research question 1: How do the grading method and the student achievement level relate to the Algebra 2 end-of-course assessment?

a. How does the grading method (standards-based grading and traditional grading) impact the results of the Algebra 2 end-of-course assessment?

b. How does the student achievement level (honors and regular) impact the results of the Algebra 2 end-of-course assessment?
c. Is there interaction in student performance on the Algebra 2 end-of-course assessment between the grading method and the student achievement level?

The results of this study found no significant difference in EOC scores between SBG classes (M = 147.62) and traditionally-graded classes (M = 147.35). Given the research and support of SBG, this finding was surprising to the teachers; both expected the SBG classes to outperform the traditional classes. Before taking the EOC exam, students reported that they felt prepared and confident in their mathematical abilities; after taking the exam, their opinions changed, and they indicated they did not feel that they understood what was being asked of them. Many felt that they performed poorly on the test.

At the time of this project, the EOC test was very new; the study was conducted during the second year of administration of the EOC exam. Educators are constantly working to understand the EOC assessment and become better equipped to prepare their students to be successful on it; however, a large gap still exists in the knowledge and skills students possess compared to what they are expected to be able to do on the EOC test. Implementation of Common Core Standards left a significant rift in the mathematical progression of students, pushing much of the content into lower grades and raising the depth of knowledge to levels many students had never been expected to demonstrate before. Additionally, Quality Core, the vendor for the EOC exam, has a separate set of standards used to dictate the questions offered on the EOC test. While they

claim their standards and the Common Core Standards are the same, there may be some discrepancy.

Relatively speaking, Common Core and Quality Core are newcomers to the educational arena. Writing learning targets that move students from where they are now to where these curriculum models want them to be is similar to shooting an arrow at an unknown target. Although the standards provide direction and guidance, there is no definite bull's eye to aim toward. Even seasoned teachers well versed in SBG may have struggled to write learning targets that aligned to the curriculum and to the EOC. Adjusting to new standards and a new testing caused difficulty in target alignment and appropriate depth of knowledge.

The second part of research question 1 asked whether there was a difference in the EOC test performance for students of different achievement levels. The analysis of the test results showed a significant difference between honors students (M = 148.10) and regular students (M = 146.84). Based on these results, honors students outperformed their regular counterparts on the EOC exam.

Students who choose to participate in honors math classes typically are students who pay more attention to detail and who have demonstrated a solid foundation in earlier math classes. Many of them enjoy math more than the students in regular level courses, and it is no surprise that they scored better on the EOC test.

Finally, a study of the results showed no significant interaction between the grading method and the achievement level. Honors students who were in the SBG classes (M = 148.12) did not do better than their regular counterparts in the traditionally-graded classes (M = 148.07). Regular students in SBG classes (M = 147.12) scored better than

their traditional counterparts (M = 146.56), but this difference was not great enough to be significant. It would appear that SBG does not have a greater impact for one achievement level over any other on the EOC exam.

While unexpected, the results of this analysis were in line with the results of Craig (2012) and the study of fourth-grade Massachusetts Comprehensive Assessment System (MCAS) results. That study showed no significant difference on math scores between SBG schools and non-SBG schools. Likewise, Hartnell (2011) did not find a statistically significant difference for students who were graded using SBG on the scores in an American government class. Together, these studies suggest that perhaps further research is necessary before SBG can be effectively implemented.

Although no significant difference was found in EOC scores for students who were members of the SBG classes, most of those students appreciated the opportunities SBG provided them. Students reported that they felt a sense of ownership of their grades, a feeling of control that they had never experienced before. The implementation of learning targets allowed them to visibly "see" everything they had learned and were able to do and presented them with a clear method of identifying the topics with which they were struggling. The students never abandoned their notion of a final grade, but they began to exhibit signs that suggested they were beginning to value mastery learning more than before. These may be the first steps in shifting their motivation to a more intrinsic one (Ryan & Deci, 2000).

Discussion of Research Question 2

Research Question 2: How do the grading method and the student achievement level impact the classroom posttest results?

a. How does the grading method (standards-based grading and traditional grading) impact the classroom posttest results?
b. How does the student achievement level (honors and regular) impact the classroom posttest results?
c. Is there an interaction in student performance on the classroom posttest results between the grading method and the student achievement level?

Students in the SBG class (M = 11.75) did not score significantly different on the post test than did the traditionally-graded students (M = 12.72). Students were administered a pretest during the first three days of the class, and the same exam was administered as a posttest during the final two days. The expectation was that the SBG students would show greater gains on the posttest score than the traditional students. When that did not happen, the researcher questioned the posttest construction. The posttest questions were taken from a set of questions released by Quality Core, the same company that authored the EOC exam. It stands to reason that, if the learning targets were not appropriately aligned to the EOC assessment, they probably would not be aligned to the posttest. This lack of significant results on this post hoc analysis highlights the effect of the limitations in effect during this research study.

Once again, honors students (M = 13.63) outperformed their regular counterparts (M = 10.94) on the posttest, as on the EOC exam. The reasons for this are largely the same as the reasons they scored better on the EOC exam; overall honors students are stronger mathematically, possess better test-taking skills, and pay greater attention to details.

No significant interaction was found between the grading method and the achievement level. Honors SBG students (M = 12.56) and honors traditionally-graded students (M = 14.65) did not score significantly different on the posttest. Also, regular SBG students (M = 10.96) and regular traditionally-graded students (M = 10.93) showed no significant difference. Interestingly, the traditionally-graded honors students did better than the SBG honors students on the posttest. Although the difference is not great enough to be significant, it introduces the notion that, perhaps, SBG can be more helpful with one particular type of student.

Sanders and Anderson (2010) reported that students' reconciliation of the meaning conveyed by a grade, especially disappointing grades, is a complex phenomenon. For students, grades are linked strongly to personal emotions, including anger, frustration, and success; they also are conditioned to believe the only successful grade is an A. In many cases, a student earning anything less than an A feels like a failure. In light of this, it would appear that standards-based grading would offer students the chance they seek, the opportunity to pursue success through reassessment. However, Sanders and Anderson (2012) also found that students were reluctant to discuss their grades with others, including their teachers; in many cases, even when students were confident enough to approach a teacher about a grade, the question was not about the degree to which the material was understood, but about "how to get a grade changed" (p. 52). This brings to the forefront a major issue with these results and SBG: students (and their parents) are firmly fixed to judge success or failure by the measure of the GPA and by extension to a final class average.

Conclusions

The final conclusions of this research project are that honors students outperform regular level students in Algebra 2. Students in the SBG classes did not outperform their peers in the traditionally-graded classes. Considering the unexpected results obtained at the conclusion of this study, a careful assessment of the limitations of this design is critical. Although this study failed to find a significant difference in SBG classes when compared to traditionally-graded classes, it can be viewed as an important source of information for teachers who are considering implementing SBG in their classrooms. What can be taken from this study is that SBG strategies require careful planning, dedication, and follow through. It is not an endeavor to be entered into lightly, but rather, with the appropriate amount of time, resources, and preparation it can provide students the chance to learn content at a mastery level. The mastery of material will, in turn, translate into higher success on high-stakes testing.

Limitations of the Study

One limitation observed by both teachers was the amount of time each had with the students in the study. The school operated on a 12-week trimester course, and the implementation of SBG was such a new concept to the students that the term was almost half over before many seemed to begin to understand exactly what was expected of them and how they could benefit from the system. It also limited the time in which students could develop mastery of the learning targets. The entire term felt rushed and frantic to both the students and the teachers.

Second, as with most initiatives, teachers need time to perfect their system of SBG. While both teachers of this study had read much literature related to learning

targets, grading, and SBG strategies, the fact that it was the first experience for them to fully implement SBG in the classroom is a consideration when interpreting these results. With practice and experience, the SBG classroom could become more finely tuned and advantageous to the students, according to much of the current research.

The researcher as one of the instructors may be viewed as a limiting factor to the study. To address this concern, weekly collaboration between the two teachers was conducted. This allowed the process to be as transparent as possible, maintaining the integrity of the process and the results. To assist with potential bias on the part of the researcher, all instruments were multiple choice exams. This mode of assessment removed any grading bias that could have been present if constructed response questions were used. Multiple choice questions remove the subjectivity that is present when scoring open-ended questions by reducing the responses to either "right" or "wrong." Additionally, the researcher did not score the posttest or the EOC questions.

The school in which the study took place is similar to many American high schools, in that each student needed a letter grade assigned to their transcript. This requirement was a definite limitation to the study; it was not possible to issue a SBG report card, and manipulating SBG procedures to produce a grade that "fits" into the traditional transcript results in the distortion of a truly standards-based learning environment. However, the main components of SBG were all present and the study can be viewed as a valid analysis of the effects of SBG versus traditional grading on student achievement.

Both instructors who participated in this study have attended professional development sessions about implementing SBG procedures, writing effective learning

targets, and employing formative assessment as a classroom tool. However, this study was the first time for either of them to fully implement SBG in the classroom.

By attempting to tweak SBG practices to fit in a school with traditional report cards and transcripts, at least a small part of the idea of learning for mastery can be lost. The almighty "average grade" trumps all other thoughts in the mind of the students and they forget that the idea is to not worry about the average, but rather to know all the skills required in the course. Instead of observing a deficiency in a specific area or learning target, students continued to check their average and decided to reassess based on that single number. Until SBG is a school-wide practice, this will continue to be a limitation of teachers attempting to implement it in single classrooms.

Future Research

This research study focused on the effects of SBG in a mathematics classroom. The pilot program of the school was conducted in only Algebra 2 classes; thus, one potential next step would be to consider SBG in other content areas. The research design of this study could guide educators in other subjects, providing a framework for procedures and implementation of other studies. Although the creation and evaluation of learning targets would vary greatly among disciplines, the fundamental idea of mastery learning and reassessment would remain constant. As SBG continues to evolve and play a more prominent role in educational reform, more research and suggestions for implementation in additional curricular areas will be necessary. If SBG is to become the norm for grading, teachers will need extensive professional development, offered over a period of years, and support as they transition to something completely new and, for many, potentially uncomfortable. Long-term consultants will be critical to its

implementation and success. Also, parents and students will need to be re-educated on what success means and the value placed on grades; they will need to learn how to decipher the new progress reports on skills and not rely as heavily on the philosophy of grade-point averages as a measure of achievement.

Two components of this design are worthy of consideration on their own merits. Learning targets were presented to both SBG and traditionally-graded classes and offered a common language, making communication between the teacher, students, and parents much more streamlined and efficient. Possibly the mere implementation of the LTs was enough to cause all students to perform better than they would have without them and offers one avenue for further research.

Students perform at higher levels when they engage in self-assessment (Marzano, 2006; Black & Wiliam, 2009), and a critical component for effective formative assessment is for students to have time to participate in self-reflection and self-assessment (O'Connor, 2009). If it is a vital component of the formative assessment and SBG processes, how can it best be accomplished? This study suggests that merely tracking performance is not enough for high school students to gauge what they need to do to improve; additional research on the most effective ways for students to self-assess and then develop a plan of action would assist other teachers who want to implement and develop a system of SBG that would work for their individual classrooms.

The results of this research suggest that a longitudinal study of SBG and its impact on standardized testing is necessary. Having students for only 12 weeks severely limited the time for building the rapport and relationship with the students that are critical for individual students to feel confident enough to request additional help and to realize

that reassessing is not a punishment, but a great opportunity. If SBG concepts could be ingrained in students at a much earlier age, education genuinely could be reformed.

REFERENCES

Alliance for Excellent Education. (2010). *High school dropouts in America.* Retrieved from www.all4ed.org

Anderson, R., & Pavan, B. (1993). *Nongradedness: Helping it to happen*. Lancaster, PA: Technomic.

Bell, T. (1993). Reflections one decade after A Nation at Risk. *Phi Delta Kappan, 74*(8), 592-597.

Biberman-Shalev, L., Sabbagh, C., Resh, N., & Kramarski, B. (2011). Grading styles and disciplinary expertise: The mediating role of the teacher's perception of the subject matter. *Teaching and Teacher Education, 27,* 831-840.

Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan International.* Retrieved from http://blog.discoveryeducation.com

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

Black, P., & Wiliam, D. (2006). Developing a theory of formative assessment. In J. Gardner (Ed.), *Assessment and Learning* (pp. 81-100). Thousand Oaks, CA: Sage.

Bloom, B. (1956). *Taxonomy of educational objectives.* White Plains, NY: Longman.

Bowers, A. (2011). What's in a grade: The multidimensional nature of what teacher-assigned grades assess in high school. *Educational Research and Evaluation: An International Journal on Theory and Practice, 17*(3), 141-149.

Brewer, E., & Kuhn, J. (2010). Causal-comparative design. In N. Salkind (Ed.), *Encyclopedia of research design* (pp. 125-132). Thousand Oaks, CA: SAGE Publications, Inc. doi: http://dx.doi.org/10.4135/9781412961288.n42

Brookhart, S. (1993). Teachers' grading practices: Meaning and values. *Journal of Educational Measurement, 30*(2), 123-142.

Brookhart, S. (1994). Teachers' grading: Practice and theory. *Applied Measurement in Education, 7*(4), 279-301.

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

Brookhart, S. (2011). Starting the conversation about grading. *Educational Leadership, 69*(3), 10-14.

Cox, K. (2011). Putting classroom grading on the table: A reform in progress. *American Secondary Education, 40*(1), 67-87.

Craig, T. (2012). *Effects of standards-based report cards on student learning* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses. (3498282)

Crooks, T. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research, 58*(4), 438-481.

Crozier, J. (2002). A unique experiment. *China in Focus,12*. Retrieved from www.sacu.org

Cureton, L. (1971). The history of grading practices. *Measurement in Education, 2*(4), 1-9.

Deddeh, H., Main, E., & Fulkerson, S. (2010). Eight steps to meaningful grading. *Phi Delta Kappan, 91*(7), 53-58.

Duncan, A. (2012, December 11). *How do U. S. students compare with their peers around the world?* Retrieved from www.huffingtonpost.com

Education Commission of the States (ECS) State Policy Database. (2007). *Standard high school graduation requirements (50 states).* Retrieved from www.ecs.org

Field, A. (2009). *Discovering statistics using SPSS.* Thousand Oaks, CA: Sage.

Goals 2000: Educate America Act. (1994). HR. 1804, 103d Cong.

Godfrey, K. (2011). *Investigating grade inflation and non-equivalence.* (Research report 2011-2). Retrieved from www.collegeboard.org

Goulden, N., & Griffin, C. (1995). The meaning of grades based on faculty and student metaphors. *Communication Education, 44,* 110-125.

Gravetter, F., & Wallnau, L. (2009). *Statistics for the behavioral sciences.* Belmont, CA: Wadsworth.

Gredler, M. (2011). Understanding Vygotsky for the classroom: Is it too late? *Educational Psychology Review, 24*, 113-131. doi: 10:1007/s10648-011-9183-6

Guskey, T. (2001). Helping standards make the grade. *Educational Leadership, 59*(1), 20-27.

Guskey, T. (2004). 0 alternatives. *Principal Leadership, 5*(2), 49-53.

Guskey, T. (2011). Stability and change in high school grades. *NAASP Bulletin, 95*(2), 85-98. doi: 10.1177/0192636511409924

Guskey, T., Jung, L., & Swan, G. (2011). Grades that mean something. *Phi Delta Kappan, 93*(2), 52-57.

Hamilton, L., Stecher, B., & Yuan, K. (2008). *Standards-based reform in the United States: History, research, and future directions* (RAND Corporate report). Retrieved from RAND corporation website: www.rand.org

Haptonstall, K. (2010). *An analysis of the correlation between standards-based, non-standards-based grading systems and achievement as measured by the Colorado StudentAassessment Program (CSAP)* (Doctoral dissertation). Available from ProQuest Dissertations and Theses. (3397087)

Hardegree, A. (2012). *Standards-based assessment and high stakes testing: Accuracy of standards-based grading* (Doctoral dissertation). Available from ProQuest Dissertations and Theses. (3540113)

Hartnell, B. (2011). *Standards-based curriculum, differentiated instruction, and end of course assessments* (Doctoral dissertation). Available from ProQuest Dissertations and Theses. (3443953)

Hattie, J. (2009). *Visible learning.* New York, NY: Routledge.

Jensen, P., & Moore, R. (2008). Do students' grades in high school biology accurately predict their grades in college biology? *Journal of College Science Teaching, 37*(3), 62-65.

Kentucky Department of Education. (2013). *School report card.* Retrieved from Kentucky Department of Education website: http://applications.education.ky.gov/SRC/Profile.aspx

Marzano, R. (2006). *Classroom assessment & grading that work.* Alexandria, VA: ASCD.

Marzano, R., & Heflebower, T. (2011). Grades that show what students know. *Educational Leadership, 69*(3), 34-39.

McMillan, J. (2001). Secondary teachers' classroom assessment and grading practices. *Educational Measurement: Issues and Practices, 20*(1), 20-32.

Moss, C., & Brookhart, S. (2012). *Learning targets.* Alexandria, VA: ASCD.

National Commission on Excellence in Education (1983). *A Nation at Risk: The imperative for educational reform.*[An open letter to the American people]. Retrieved from www.2.ed.gov

National Governor's Association Center for Best Practices, Council of Chief State School Officers. (2012). *Common Standards.* Retrieved from www.corestandards.org

O'Connor, K. (2009). *How to grade for learning.* Thousand Oaks, CA: Corwin.

O'Connor, K. (2011). *A repair kit for grading: 15 fixes for broken grades.* Boston, MA: Pearson.

O'Shea, M. (2005). *From standards to success.* Alexandria, VA: ASCD.

Pedhazur, E., & Schmelkin, L. (1991). *Measurement, design, and analysis.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Quality Core. (2010). *Technical manual.* Retrieved from www.act.org

Quality Core. (2012). *Algebra II EOC test blueprint.* Retrieved from www.act.org

Rudalevige, A. (2003). The politics of No Child Left Behind. *Education Next, 3*(4), 63-69.

Ryan, R., & Deci, E. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology, 25,* 54-67. doi: 10.1006/ceps.1999.1020

Sadler, D. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18,* 119-144.

Sanders, M., & Anderson, S. (2010). The dilemma of grades: Reconciling disappointing grades with feelings of personal success. *Qualitative Research Reports in Communication, 11*(1), 51-56.

Shim, S., & Ryan, A. (2005). Changes in self-efficacy, challenge avoidance, and intrinsic value in response to grades: The role of achievement grades. *The Journal of Experimental Education, 73*(4), 333-349.

Stiggins, R. (2005). From formative assessment to assessment for learning: A path to success in standards-based schools. *Phi Delta Kappan, 87*(4), 324-328.

Stiggins, R. (2007). Assessment through the student's eyes. *Educational Leadership, 64*(8), 22-26.

Stones, E. (2012). *Educational objectives and the teaching of educational psychology.* New York, NY: Routledge.

Strauss, V. (2012, September 25). President Obama interview for education nation—transcript. *The Washington Post.* Retrieved from www.washingtonpost.com

Tassell, J., Kemp, J., Litkenhus, D., & Schriefer, M. (2006). Progress report vs. report card: One district's challenge. *Technos 3*(10). Retrieved from www.ait.net

Tierney, R., Simon, M., & Charland, J. (2011). Being fair: Teachers' interpretations of principles of standards-based grading. *The Educational Forum,. 75,* 210-27. doi:10.1080/00131725.2011.577669

U.S. Department of Education. (2012). *Race to the Top assessment program.* Retrieved from www.ed.gov

Ziomek, R., & Svec, J. (1995). *High school grades and achievement: Evidence of grade inflation.* ACT research report series, 95-3.

Zirkel, P. (2007, March 28). Grade inflation: High schools' skeleton in the closet. *Education Week, 26*(29), 40, 30.

APPENDIX A: IRB Approval

A LEADING AMERICAN UNIVERSITY WITH
INTERNATIONAL REACH

OFFICE OF COMPLIANCE

DATE:	January 11, 2013
TO:	Rachel Rosales
FROM:	Western Kentucky University (WKU) IRB
PROJECT TITLE:	[414936-1] The effects of standards-based grading on student performance in algebra 2 (v.2)
REFERENCE #:	IRB 13-161
SUBMISSION TYPE:	New Project
ACTION:	APPROVED
APPROVAL DATE:	January 11, 2013
EXPIRATION DATE:	November 30, 2013
REVIEW TYPE:	Expedited Review

Thank you for your submission of New Project materials for this project. The Western Kentucky University (WKU) IRB has APPROVED your submission. This approval is based on an appropriate risk/benefit
ratio and a project design wherein the risks have been minimized. All research must be conducted in accordance with this approved submission.

This submission has received Expedited Review based on the applicable federal regulation.

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

All UNANTICIPATED PROBLEMS involving risks to subjects or others and SERIOUS and

UNEXPECTED adverse events must be reported promptly to this office. Please use the appropriate reporting forms for this procedure. All FDA and sponsor reporting requirements should also be followed.

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

This project has been determined to be a Minimal Risk project. Based on the risks, this project requires continuing review by this committee on an annual basis. Please use the appropriate forms for this procedure. Your documentation for continuing review must be received with sufficient time for review and continued approval before the expiration date of November 30, 2013.

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

If you have any questions, please contact Paul Mooney at (270) 745-2129 or irb@wku.edu. Please include your project title and reference number in all correspondence with this committee.

This letter has been electronically signed in accordance with all applicable regulations, and a copy is retained within Western Kentucky University (WKU) IRB's records.

APPENDIX B: DCPS Letter of Support

Daviess County Public Schools

1622 Southeastern Parkway
Owensboro, KY 42303

Owens Saylor
Superintendent

Daviess County
Board of Education
Frank G. Riney, III,
Chairperson
Dianne B. Mackey,
Vice-Chair
Merritt Bates-Thomas
Randall E. King, M.D.
P. Mike Clark

www.daviess.kyschools.us
270-852-7000 (phone)
270-852-7030 (fax)

December 26, 2012

To Whom It May Concern:

This letter is to serve as support for Rachel Rosales' research study on the effects of standards-based grading on student performance in algebra 2. The district fully supports her planned study of four sections of algebra 2, effective February 2013-May 2013 and the use of student test scores for analysis of her research.

Daviess County Public Schools is refocusing its vision on student achievement and performance, working to build Great Expectations for all of our students. One component of that initiative is the creation and use of data systems that measure student growth. Standards-based grading provides a mechanism for that measurement at the classroom level. In addition, standards-based grading helps serve a second part of our focus, helping students assume more responsibility for their own learning.

Additionally, an important element to our districts commitment as a recipient of Race to the Top Federal Grant dollars involves a shift to personalized, competency-based instructional practices, designed to assist students in authentically mastering standards which is at the very core of standards-based grading.

Standards-based grading is emerging as an important initiative across the nation and as a strong proponent of the work of Dr. Thomas Guskey, I am eager to see it blossom in our district. Mrs. Rosales has the full support of the administrative team of her school and the district office. We eagerly await the findings of her research.

Sincerely,

Owens Saylor, Superintendent
Daviess County Public Schools

APPENDIX C: Letter to parents of students in traditionally-graded classes

Dear parents:

Welcome!

I am excited to be your child's Algebra 2 teacher this term. I hope they are ready for a fun-filled term, full of learning.

Before the term begins, I would like to share a couple of items that will help your child enjoy the maximum amount of success. First, mathematics is a very difficult subject to learn independently. It is vital that your student is in attendance every day, fully prepared to participate in the activities for that day. Fully prepared means coming to class with all required materials (completed homework assignments, notebook, colored pencils, graphing calculator, and a charged computer) and coming to class with the intent to focus and work hard. Sleeping and lazy attitudes will most definitely hinder success.

Second, I am available to help your child when they need it. I am available on Mondays and Wednesdays from 7:40-8:10 and Mondays, Wednesdays, and Thursdays from 3:15-3:45. Please encourage your student to see me during these times if they find material difficult. During our class there is a lot of activity and we move quickly; many students find that they need a small amount of one-on-one time to review skills or topics and I am more than happy to meet with them. If they cannot come during the times listed, have them contact me to set an appointment for a different day.

Finally, the largest percentage of your child's grade will be based on work completed in class. Homework is critical for practicing skills and will count for 10% of the overall term grade. In class work offers your child an opportunity to demonstrate proficiency and will count as 90% of the overall term grade. During each unit students will receive a list of the learning targets (LT) required to successfully pass the unit.

If you have any questions, please feel free to contact me.

Thank you.

APPENDIX D: Letter to parents of students in SBG classes

Dear parents:

Welcome!

I am excited to be your child's Algebra 2 teacher this term. I hope they are ready for a fun-filled term, full of learning.

Before the term begins, I would like to share a couple of items that will help your child enjoy the maximum amount of success. First, mathematics is a very difficult subject to learn independently. It is vital that your student is in attendance every day, fully prepared to participate in the activities for that day. Fully prepared means coming to class with all required materials (completed homework assignments, notebook, colored pencils, graphing calculator, and a charged computer) and coming to class with the intent to focus and work hard. Sleeping and lazy attitudes will most definitely hinder success.

Second, I am available to help your child when they need it. I am available on Mondays and Wednesdays from 7:40-8:10 and Mondays, Wednesdays, and Thursdays from 3:15-3:45. Please encourage your student to see me during these times if they find material difficult. During our class there is a lot of activity and we move quickly; many students find that they need a small amount of one-on-one time to review skills or topics and I am more than happy to meet with them. If they cannot come during the times listed, have them contact me to set an appointment for a different day.

Finally, I will be using a grading system that allows students to retest until they reach proficiency or mastery on individual goals. At the beginning of each unit students will receive a list of the learning targets (LT) required to successfully pass the unit. Learning targets will be graded individually using the attached rubric.

Students can revise their scores on each learning target during the term. To improve their grade, students will be required to practice problems related to the target, and take a new quiz on that target. It is important that students know they cannot sign up to re-assess on the same day that they seek extra help for practicing the standard. This is to prevent students from memorizing a set of steps for a short period of time. The intent is to encourage long-term retention.

If you have any questions, please feel free to contact myself or Kyle Brown, assistant principle. Thank you.

APPENDIX E: Learning Targets for regular level students

Regular Level Learning Targets

Learning target	CCS Standard	Quality Core Standard
F1: I can evaluate function operations numerically and algebraically.	F-IF 1 F-IF 2	C.1d
F2: I can evaluate composite functions numerically and algebraically.	F-BF 1b	C.1.d
F3: I can solve equations using function operations.	A-REI 3	C.1.d
P1: I can determine zeros of polynomial functions.	A-APR 3	F.2.a F.2.b F.2.c
P2: I can determine end behavior of polynomial functions.	F-IF 4	F.2.d
P3: I can determine the behavior at the zeros of a polynomial function.	F-IF 4	F.2.d
P4: I can write multiple equations using zeros, end behavior, and behavior of zeros.	F-IF 4	F.2.c
P5a: I can convert from standard form to factored form by the GCF method.	F-IF 8	F.1.b
P5b: I can convert from standard form to factored form by synthetic division.	F-IF 8	F.1.b
P6: I can divide polynomials with long division.	A-APR 6	F.1.b
R1: I can convert to a fractional exponent.	N-RN 2	G.1.f
R2: I can solve equations dealing with radical and exponents.	A-REI 2	G.1.b G.1.f
R3a: I can simplify radicals involving letters.		G.1.b G.1.c
R3b: I can simplify radicals involving numbers.		G.1.b G.1.c G.1.d
E1a: I can find the following from an exponential model: rate (%).	F-LE 5	
E1b: I can find the following from an exponential model: asymptote.	F-IF 4 F-LE 5	
E1c: I can find the following from an exponential model: y-intercept.	F-IF 4	
E2: I can write an exponential growth model and answer questions using that model.	A-CED 2 F-BF 1a F-LE 1c F-LE 2	

E3: I can write an exponential decay model and answer questions using that model.	A-CED 2 F-BF 1a F-LE 1c F-LE 2
E4: I can use $A = P\left(1 + \frac{r}{n}\right)^{nt}$ to find the amount of money with compounded interest.	F-BF 1a
E5: I can find an inverse of a function using a variety of operations.	F-BF 4

APPENDIX F: Learning Targets for honors level students

Honors level learning targets

Learning target	CCS standard	Quality Core Standard
1a: I can evaluate functions using tables.	F-IF 1 F-IF 2	C.1d
1b: I can evaluate functions using graphs.	F-IF 1 F-IF 2	C.1.d
1c: I can evaluate functions using equations.	F-IF 1 F-IF 2	C.1.d
2: I can combine functions using addition, subtraction, multiplication, and composition.	F-BF 1b F-BF 1c	C.1.d
3: I can graph a polynomial function.	A-APR 3	F.2.d
4: I can factor higher-order polynomials.	A-SSE 3a	F.1.b
5: I can find all the roots of a polynomial functions (factoring, quadratic formula, synthetic division).	A-APR 2 A-APR 3	F.2.a F.2.b F.2.c
6: I can divide polynomials (synthetic and long division).	A-APR 6	F.1.b
7: I can write the equation of a circle.	G-GPE 1	E.3.d
8: I can solve equations containing radicals and exponents.	A-REI 2	G.1.f
9: I can simplify radicals.		G.1.b G.1.c
10: I can write and evaluate equations for exponential functions.	A-CED 2 F-BF 1a F-LE 1c F-LE 2	
11: I can find inverse functions using graphs, tables, and equations.	F-BF 4	
12: I can evaluate logarithms.	F-BF 5	G.2.b
13: I can solve exponential and logarithmic equations algebraically.	F-BF 5 F-LE 4	
14: I can determine the period and amplitude of a sine and cosine graph.	F-TF 5	G.3.f
15: I can convert between degrees and radians.	F-TF 1	G.3.c