THE EFFECTS OF STANDARDS-BASED GRADING IN A SECONDARY MATHEMATICS CLASSROOM

A doctoral thesis presented by
Joshua Martin Bromley
to
The Graduate School of Education

In partial fulfillment of the requirements for the degree of
Doctor of Education
in the field of
Education

College of Professional Studies
Northeastern University
Boston, Massachusetts
March 2019
Abstract

The current generation of American students find themselves ill-prepared for the post-secondary level of mathematics study due to graduating high school without the requisite skills to be successful. While there are many factors that contribute to this deficit, it remains the job of educators and educational researchers to make strides toward bridging that gap. While standards-based pedagogical and curricular reforms have increased in popularity and predominance, this study aimed to examine the impact of student grading and evaluation methodology on overall student achievement. Standards-based grading, a long-time practice in elementary schools, was implemented in two secondary mathematics classrooms at a rural school in western North Carolina. Through the use of detailed rubrics and feedback tracking worksheets, students were exposed to this method of grading throughout an entire semester in their Math 1 classes. To understand how these students would differ from traditionally-graded students, a covariate analysis was carried out to determine the significance of grading systems upon student achievement on the North Carolina End-of-Course exam for Math 1. This analysis subsequently showed a significant relationship between membership in the standards-based classes with students showing higher levels of achievement than their traditional counterparts. Additionally, focus group interviews were held with teachers and students, both of whom gave positive feedback regarding the standards-based system of grading, attesting to the specificity of the evaluation rubrics and the increased level of preparedness.

*Keywords*: Standards-based Grading; mathematics; high school; evaluation; rubrics; student feedback; social cognitive theory; instructional techniques.
Acknowledgements

This is, without a doubt, the most challenging thing I have ever gotten myself into, and this is after raising three children and being a high school administrator. I am eternally grateful to all those who have helped me as I have travelled down this path.

Dr. Michael Dean, thank you for being patient with me and helping me persevere through obstacles and difficulties. Your expertise and guidance through a world with which I was not familiar greatly benefited my research and sanity. I would also like to thank the other members of my committee, Dr. Christopher Unger and Dr. Thomas Potter, for their time and valuable feedback throughout this process.

I would also like to thank Brian Clary, my former principal at West Lincoln High School for being receptive to the questions and ideas proposed by this study. Without his flexibility in scheduling and understanding, this study would not have happened. Additionally, I owe a large debt of gratitude to the math teachers participating in this study for letting me turn their world upside down for an entire semester.

Very few people in my life truly understand the time and work I have put into my studies, but my wife Jasmine is one of those people. Countless times when I had to work late, work early, or lock myself in a room for long periods of time to focus, she took care of the kids, the house, and anything else that came up. Jasmine, thank you all you have done and continue to do.

Lastly, My parents, JoAnn and Charles, and my grandmother, Tina, have always been supportive in my education, and without their encouragement and contributions, I certainly would not have ended up where I am today. They’ve supported me in everything I do and continue to do so to this day. I wouldn’t be here without you all. I love you all, and thank you so much.
Dedication

For Austin, Evan, and Isaac:
All our dreams can come true, if we have the courage to pursue them.
# Table of Contents

Abstract .................................................................................................................. 2  
Acknowledgement ................................................................................................. 3  
Dedication .............................................................................................................. 4  
Table of Contents ................................................................................................ 5  
List of Tables ....................................................................................................... 7  
List of Figures ..................................................................................................... 8  
Chapter 1 ............................................................................................................. 9  
  Problem Statement .............................................................................................. 9  
  Problem Significance ......................................................................................... 11  
  Positionality ...................................................................................................... 12  
  Purpose Statement ............................................................................................ 15  
  Research Questions ......................................................................................... 15  
  Theoretical Framework .................................................................................... 16  
    Analysis of Theory ........................................................................................ 18  
      Rewards and punishment ......................................................................... 20  
      Human observation ................................................................................... 20  
      Individual behavior .................................................................................. 20  
Chapter 2 ........................................................................................................... 22  
  Introduction ..................................................................................................... 22  
  Student Performance in Mathematics ............................................................. 25  
  Traditional Grading Systems ....................................................................... 28  
  Standards-based Grading Systems ................................................................. 31  
  Advocacy Argument ....................................................................................... 35  
  Conclusion ....................................................................................................... 37  
Chapter 3 ........................................................................................................... 39  
  Research Questions ......................................................................................... 39  
  Purpose of the Questions ............................................................................... 39  
  Discussion of the Variables .......................................................................... 39  
  Hypotheses / Rationale .................................................................................. 40  
  Research Design .............................................................................................. 40  
  Justification for the Design ......................................................................... 41  
  Implications ..................................................................................................... 41  
  Limitations for the Design ......................................................................... 41  
  Site and Participants ...................................................................................... 42  
  Protection of Human Subjects .................................................................... 44  
  Instruments ....................................................................................................... 44  
  Procedures ....................................................................................................... 46  
  Validity, Reliability, and Generalizability ...................................................... 48  
Chapter 4 ........................................................................................................... 50  
  Introduction ..................................................................................................... 50  
  Quantitative Results ....................................................................................... 50  
    Pre-test and Post-test Scores .................................................................... 50  
  Qualitative Results ......................................................................................... 57  
  Summary of Quantitative and Qualitative Data ........................................... 60  
Chapter 5 ........................................................................................................... 62
# List of Tables

<table>
<thead>
<tr>
<th>Tables</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Student Pretest Results Control</td>
<td>51</td>
</tr>
<tr>
<td>2. Student Pretest Results Experimental</td>
<td>51</td>
</tr>
<tr>
<td>3. Correlation of Covariate</td>
<td>51</td>
</tr>
<tr>
<td>4. Student Postest Results Control</td>
<td>52</td>
</tr>
<tr>
<td>5. Student Postest Results Experimental</td>
<td>52</td>
</tr>
<tr>
<td>6. Student Growth Control</td>
<td>53</td>
</tr>
<tr>
<td>7. Student Growth Experimental</td>
<td>53</td>
</tr>
<tr>
<td>8. Group and Pretest Interaction</td>
<td>54</td>
</tr>
<tr>
<td>9. Significance of Group Membership</td>
<td>54</td>
</tr>
<tr>
<td>10. Analysis of Variance</td>
<td>55</td>
</tr>
<tr>
<td>11. Coefficients of Regression</td>
<td>55</td>
</tr>
<tr>
<td>12. Group and Teacher Interaction</td>
<td>56</td>
</tr>
<tr>
<td>13. Effect of Group Membership on Postest by Teacher</td>
<td>56</td>
</tr>
<tr>
<td>14. Thematic Analysis of Student Interview Responses</td>
<td>58</td>
</tr>
<tr>
<td>15. Thematic Analysis of Teacher Interview Responses</td>
<td>59</td>
</tr>
</tbody>
</table>
List of Figures

Figures                                      Page
  1. Mean Achievement Variance of Postest Scores by Teacher ....................... 57
American high school students often graduate high school without the requisite skills for assimilation into the job market or society as a whole. As a result, they are forced to enroll in remedial courses upon entering post-secondary studies, and as research shows, are often less successful in their overall courses of study as a result. Examining the problem from a global perspective reveals a picture painted with substandard mathematics scores when compared with their counterparts in other countries.

With the widening mathematics achievement gap between American students and their foreign counterparts, it is not surprising that a breadth of research has been devoted to improvements in this area. Typically research is based in pedagogical or curricular reforms, citing best practices and new innovative teaching techniques to deliver a more rigorous and practical curriculum that emphasizes critical thinking and problem solving (Lee & Huh, 2014; Graber, 2006, Brantlinger, 2013; Marzano & Gaddy, 2000). While these may be the most obvious influences upon classroom instruction and achievement, very rarely has the lens been shifted to the ways in which students are evaluated – namely the grading system employed to measure student mastery. Student assessment methodologies and grading paradigms are a becoming an increasingly popular topic of exploration as those in the field of education seek to find ways to optimize academics and improve student achievement. Relevant research has shown that in addition to the way students are taught, the way in which students are evaluated for mastery can have an impact on their perceived and actual achievement (Togolini & Stanley, 2007; Silver & Kenney, 1995).
Traditional grading systems often employ a simplistic rubric painted with broad strokes that label students as average, above average, or below average, and subsequently assign them a letter grade between A and F to reflect overall achievement (Lekholm & Cliffordson, 2008). This grading system, while beneficial in its ease of adoption and use, ultimately fails students by not providing the specific feedback necessary to better clarify their educational strengths and weaknesses. It is a creature of comfort, and exists out of familiarity and pre-determined widespread implementation (Allen, 2005; Ausin & McCann, 1992). This research will focus upon a grading alternative, namely standards-based grading, and the effect of its adoption and use in the secondary mathematics classroom as a function of overall student achievement on standardized mathematics assessments.

Mathematics teachers strive for their students to be highly literate in the content delivered to them in the classroom. Successful students are bred in successful classrooms that exhibit the constructs of an environment that is conducive and supportive to student achievement. With that in mind, everything that takes place in the classroom should be closely examined when exploring avenues for increasing student achievement. Pedagogy is typically the component first examined—with good reason—but this examination of research supports evidence that educators need to sometimes look beyond that. Policies and procedures dictating the ways students are evaluated in the classroom, in both a formative and summative means, can influence the ways students achieve, and perceive their achievement. These achievements then correlate directly with educational attainment.

The problem presented herein is that student achievement in the mathematics classroom suffers negatively from a traditional percentage-based grading schema. In turn, students are ill prepared for sequential post-secondary studies requiring mathematical proficiency. To gain a
better understanding of the existing body of literature, this review will explore educational research regarding mathematics achievement, traditional grading schema, and standards-based grading schema.

The research will culminate in the implementation of a mixed-methods approach in which student test data is examined along focus group narratives, a more clear determination will be made of the benefits of alternative grading systems and their effect on student achievement as a product of increase morale, clarity, stakeholder involvement, and enthusiasm for learning.

**Problem Significance**

The importance of mathematic literacy with regard to post-secondary success can not be overstated. Multiple research studies show that students that are proficient at mathematics and in turn complete a more rigorous course of study are much more likely to have substantially higher earning potential (Rose & Betts, 2001) upon exiting high school. Science, Technology, Engineering and Mathematics (STEM) based careers are growing at four times the rate of total employment over the past twenty years (Hossain & Robinson, 2012) and qualified individuals are needed to fill those roles. With a deficit in US graduates, these high paying stem careers are typically filled by foreign-born workers (Hossain & Robinson, 2012) that have learned the requisite skills at foreign educational institutions.

The research is clear that United States students demonstrate a fundamental lack of basic mathematics skills upon entering high school (Silver, 2014) with a large majority of students failing to perform on grade level. The National Assessment of Educational Progress (NAEP) paints a picture of student performance in both reading and mathematics in fourth and eighth grade across the United States. According to data from the 2013 Nation’s Report Card, 64% of eighth grade students score less than proficient in mathematics (National Center for Education
Statistics, 2013). Students that are not proficient typically end up in some form of remedial math class that reinforces basic skills (Brantlinger, 2013) in an effort to better prepare them for future math classes, and in turn give them skills needed to graduate and subsequently enter the workforce. Research suggests that upon graduation, a majority of students still lack basic skills with a majority requiring remedial assistance in their mathematics studies (Bahr, 2013) to be successful at a post-secondary level.

When examining the performance of United States students with respect to the rest of the world, students are also lacking in performance. Students in the United States score on average 75 – 100 points lower than students from countries with similar GDP (Lee & Huh, 2014). This deficiency in mathematics ability produces students that will have difficulty competing in the globally competitive marketplace (Silver, 2014) that exists today.

Students in North Carolina, when compared to their counterparts in the rest of the United States, are admittedly average performers (National Center for Education Statistics, 2013). North Carolina ranks 26th in the nation on the eighth grade NAEP test, with 36% of students testing as proficient – an identical figure to that of the country as a whole. Viewed from any angle, the significance of this problem is far better supported by a situation in which only a third of participants are successful warrants cause for some sort of drive for improvement.

**Positionality Statement**

My problem of practice is the effect of traditional and alternative grading schema upon student achievement in secondary mathematics. As a mathematics teacher at the secondary level for the past decade, I have witnessed first hand the benefits and consequences brought to the classroom by the current educational mindset when it comes to grading students, and the direct effects it can have on student motivation and overall achievement. Student disillusionment and
misrepresentations of student mastery have motivated me to pursue alternative and less conventional grading frameworks. In order to do this in a subjective way, it is important that I consider my positionality with regard to the problem of practice I wish to investigate further.

Positionality is a complex and diversified view of one’s relationship to a situation that is obtained by examining several facets of one’s experiences and environment. “The ideas that individuals construct an understanding of the world and perceive themselves to occupy a particular location within the reality they construe are key premises to positionality (Kinchetlow & Steinburg, 1998, pg. 21). By looking at my experience over my career as an educator, I understand that I bring with me my own viewpoints, understandings, and predetermined biases. I believe our current grading system, in which every student is assigned a letter grade based on a percentage of overall mastery does many a disservice by labeling them failures when they may have experience varying levels of success in certain concepts or applicational skills. I think that as a result, students become disinterested, unmotivated, and enter a nearly irreversible downward spiral of shut-down that permeates throughout the entire semester, and can similarly be witnessed throughout that students educational path. From this, students develop mindsets that they’ve “never been good at math” or their family has “never been good at math” – anything to really find an easily attributable diagnosis for the problem. In my opinion, the grading system currently in place in schools across the country exists for the convenience of teachers and administrators who seek to reduce every student to a numerical value, thus making it easier to compare and contrast.

My viewpoints and observations stem from several different factors that make up my positionality. “Dimensions of positionality include one’s demographic positioning within
society, one’s ideological positioning, and how one dicursively positions the other and oneself (Briscoe, 2005, pg. 28).

I am a thirty-five year old white male living outside of Charlotte, North Carolina. I am married and have three sons – a ten year-old and four year-old who attend public school, and a one week old that we welcomed into our family last week. I have been teaching at the same high school for 14 years, and have taught various levels of mathematics during that time to student of different levels of ability, age, and experience. I have had the opportunity to assume leadership roles in both my school and county, and I am heavily involved in the acquisition and implementation of educational technology at our school.

My educational background includes undergraduate studies in Computer Science and Mathematics, graduate studies in Educational technology and Executive School Leadership, and the doctoral program in which I am currently enrolled.

My family resides in Pittsburgh, Pennsylvania in the neighborhood in which I grew up as a child. Our family was predominantly middle class, both of my parents never having attended college. My father works in an industry as a skilled tradesman, and my mother mostly stayed home and cared for the myself and the house. The dedication and perserverence they have shown in their personal and professional lives has molded my work ethic in a similar fashion.

I’ve had the unique opportunity over the years to interact with youth outside of the school during my summer job as a Boy Scout summer camp director. My interactions and experiences with youth have furthered my beliefs that strengthened my opinions that every child, no matter of background, skillset, or upbringing, has something to contribute that at which they truly excel. It is the job of educators to foster that growth, and bring out the best in our students by finding a way to incorporate that into what we teach on a daily basis.
The students I teach at my school live in an extremely rural area. Many come from an agricultural background, and the percentage of parents with a college degree is well under 10% as shown in recent parent survey data. With that in mind, it could be argued that current school policy and curriculum are not conducive to the majority of learners at our school. According to Jupp, “schooling reproduces the ideologies of the groups or classes that actually rule society” (2006, pg. 201). Students at my school are often forced into advanced coursework not relevant to their career choices or pathway they have chosen to fulfill graduation requirements.

By building relationships with my students and striving to understand their goals and the happenings in their daily lives, I try to put myself in a situation in which I can conduct myself, professionally and personally, in a way that puts their best interest at the forefront. My classroom and school is truly a “complex and contested terrain” in that the participants exist in a “dicursive context” (Jupp, 2006, pg. 203) and it is important to realize that when trying to make sense and draw conclusions as such. Cole (1995) reiterates the importance of viewing the subjects of any study “in relation to institutional arrangements or contexts” (pg. 29). It is my hope that through additional analysis and scrutiny, I can approach this problem of practice in a neutral and uncompromised fashion to help understand ways in which we can make changes to increase student achievement and motivation in the classroom.

**Purpose Statement**

The purpose of this study is to compare the overall student achievement of students enrolled in a secondary mathematics course utilizing a traditional grading system and students enrolled in the same course utilizing a standards based grading system in terms of End-of-Course exam grades at West Lincoln High School in Lincolnton, North Carolina.
Research Questions

1. Does the way in which students are graded exhibit a measurable influence upon overall student achievement in a secondary mathematics class?
   
   a. Does student achievement in secondary mathematics benefit/suffer from a traditional omnibus grading system?
   
   b. Does student achievement in secondary mathematics benefit/suffer from a standards based grading system?

2. To what extent do students perceive grading systems as having influence on their achievement in secondary mathematics?

Theoretical Framework

Student achievement plays a central role in a wide variety of problems facing schools throughout the country. Gaps exist between genders, ethnicities, and socioeconomic groups that stem from social and cultural issues. Percolating a plethora of equity and justice tinted layers onto an already complex situation, makes it even more difficult to arrive at a solution. The issue remains, however, that in a globally competitive world in which requisite skills are necessary to be a viable participant, American students are falling behind.

Lack of achievement in secondary studies sets students on an often-uncorrectable course toward remedial studies, disinterest in learning, and a poor attitude toward schooling in general. Students who fail to experience success in school find themselves pigeonholed into careers with lower pay, less benefits, and enrollment in a cycle that often follows generation to generation. Therefore, it remains that increasing student achievement in school, across the board, is in the best interest of everyone involved.
The researcher’s problem of practice is the negative effect that traditional grading systems have upon students as shown in overall achievement in secondary mathematics classrooms. While much focus has been placed upon curricular reforms and the implementation of support structures through remediation and mentoring, it is the researcher’s belief that measurement of mastery and the subsequent communication of that measurement has very much fallen by the wayside. Grading reform should very much be on the table when it comes to diagnosing the problems that ail our students, because it plays a large role in the establishment of a student’s self-efficacy. Often, students develop pictures of their abilities based upon the feedback that is given to them. If teacher feedback does not truly communicate mastery, or is distorted in a fun-house mirror-like fashion, the feedback will almost certainly do more damage than good.

This study will focus upon the impact of grading systems on students in a secondary Mathematics class in rural North Carolina. Algebra classes with similar students, teachers, and curriculum, will be evaluated in two different ways throughout the semester – one employing the traditional grading methods with which we are all very familiar, and one employing a standards-based grading approach in which measurement of achievement is aligned with content standards and reported as such. Through pre- and post-semester assessments, the researcher hopes to gain a better understanding of student evaluation paradigms and their effect on student achievement. The researcher’s hypothesis is that more meaningful evaluations, coupled with the disappearance of “failing” grades, will help boost students’ self-esteem towards learning and their own abilities. This would hopefully translate to larger achievement gains when measured using the pre- and post-semester instruments.
At West Lincoln High School, mathematics achievement plays an important factor in our overall school’s effectiveness rating. Similarly, at the end of our last accreditation cycle, parental and community involvement were identified as target areas in which we could show improvement as we work toward future accreditations. Providing parents and students with an opportunity to participate in standards-based measurement helps boost those avenues of communication and involvement necessary to the stakeholders involved in the educational process. It is the researcher’s hope that detailed reports aligned with content standards will provide more useful information that can inform decisions to guide students, parents, and teachers in their learning and teaching.

Any increase seen in student achievement at West Lincoln High School paves the way for future work regarding the way in which students are evaluated in other content areas and at other grade levels. With student achievement already at the forefront of educational research, it follows logically that any type of reform, whether it be rooted in social justice, curriculum, or evaluation, that could possibly help increase student achievement, is automatically germane to any discussion. Increased levels of student achievement translate directly to benefit everyone involved in education, both directly and indirectly. If the research shows significant impact upon achievement as a product of the way in which students are evaluated, it creates traction for it to overcome more traditional systems that have stuck around for the sake of convenience and familiarity. It is already widely agreed upon that a student’s self-efficacy plays a large role in their success in school. Acknowledging that social-cognitive aspect of a student’s learning behavior helps shift from an ultimate evaluation model, to a continuous one.
Analysis of Theory

Largely researched and expanded upon by Albert Bandura (1993), Social Cognitive Theory finds its roots in early studies of behaviorism by Watson (1919) and Thorndike (1911). Seeking to better understand the ways in which learners acquire knowledge, behaviorists characterized the learning process as a trial and error exercise, driven solely by the perceived outcomes of activities. On an experimental basis, it was their view that learners gained knowledge by simply finding what works best in a situation. Bandura (1965) saw a problem with this, and argued that social constructs, as well as interactions and observations with others also played a large role in the acquisition of knowledge. Rightly so, he found it difficult to believe that the world in which we live was comprised of a culture in which its “language, morals, familial customs and practices, occupational competencies, and politics” (p. 10) were shaped entirely by trial and error performances.

Bandura conducted extensive research throughout his career (1965, 1971) that led to the conclusion that “generalized imitation is governed by social beliefs and outcome expectations” (p. 11) much more than reinforced behavioral feedback. The subjects in his studies reacted more favorably to social behaviors and perceived norms when deciding how to navigate through presented scenarios. The perceptions of others, and what we gain to inform our own perceptions, largely shapes how we conduct ourselves.

Self-efficacy, or the perception of our own abilities in certain situations, are central to Bandura’s Social Cognitive Theory. Personality, motivation, and one’s actions or reactions to external stimuli are influenced by the way one observed the external factors at play during any given time. Bandura (1993) concluded through his research that self-efficacy is largely developed in a social way, through both observational and social experience with others –
ultimately hypothesizing that those possessing a higher level of self-efficacy are more likely to persevere in difficult situations rather than giving up altogether.

Critics of Social Cognitive Theory argue that much like behaviorism, it is too unilateral to accurately describe the rich, organic experience of acquiring knowledge. Myers (2010) contends that focusing on only the social aspect of learning, observation, and other’s expectations in a situation, fails to adequately acknowledge the traits and qualities possessed by the learner – which he argues play an equally large role in a situation. However, it can be argued that Bandura accounted for these in his proposition of reciprocal determinism (1978) in which three factors – environment, individual, and behavior – all contribute to the formation of the learner and their abilities.

Social Cognitive Theory is a prevalent framework in grounding many educational research studies. Lent, Brown, Brenner, Chopra, Davis, Talleyrand and Suthakaran (2001) examined the effect of external environmental barriers on the choices students made in their mathematics and science studies. Similar studies have examined the gender gap (Spencer, Steele, Quinn, 1999) and racial gap (Wang, 2013) in STEM related fields. Additionally, Wentzel (1998) framed an argument in which the social relationships formed by students, teachers, and their peers, play a large role on the establishment of one’s self-efficacy and subsequent motivation in a middle school environment. All of these studies have an underlying theme, which is to be expected when grounded in the same theoretical framework – that social interactions and observations establish norms and expectations for students against which they measure themselves. This is the foundation of how students determine self-worth. Brown (1999) outlines three tenets for Bandura’s Social Cognitive Theory, all of which have a role in the development of this study.
Rewards and punishments increase the likelihood that a behavior will be repeated.

Feedback given to students on their performance in a class is a form of reward or punishment. Traditional grades, in the form of letters A – F, classify students as either a great success, or a failure to succeed. Students who are told they are a success get reinforced by that behavior, much like students who are told they are a failure. By eliminating these letter grades as well as the finality of assessment in lieu of a more descriptive and continuing practice, it is the researcher’s hope that the role of the internal locus will be minimized.

Human observation

Humans can learn by observing others. Students’ self-efficacy is developed largely in part by the way and manner the social environment in school functions around them. Vicarious learning, through observation, about one’s abilities and ultimately the capacity to build upon those abilities through hard work and perseverance is shaped by a student’s interactions with his peers and his teacher. It is the researcher’s hope that standards-based grading helps to create an environment in which students thrive in their interactions with peers and grow from it.

Individual modeling

Individuals are most likely to model behavior observed by those with which they identify. An environment of mutual respect between students and their teachers is fostered by interactions that show compassion, care, and genuine interest in students’ lives. Standards-based assessment removes the stigma of failing grades and missed targets instead replacing them with road maps to success. Similarly, the almost gamification of learning, in which students work together to solve problems and build each other up helps to create an environment where learners identify with each other rather than compete for GPA points.
Chapter 2: Review of the Literature

**Introduction**

A large number of students entering higher education find themselves unable to perform at a level of mathematical proficiency necessary to be successful in their schoolwork. Not only does this place students in a precarious academic situation with regard to their intended path of study, but it also solidly entrenches the majority of students well behind their international counterparts. Although the problem may become apparent at the post-secondary level, its cause is rooted in students’ prior academic studies. This deficit in skills contributes directly to the failure of students endeavoring to complete post-secondary programs as they become frustrated and ultimately drop out.

A large vacuum exists in science, technology, engineering, and mathematics (STEM) career fields that needs to be filled to prevent the industry from imploding upon itself. Curricular progress has been made in secondary schools across the country that exposes students to higher level mathematics and science. Coupled with increased academic standards for mastery and rigor in the classroom, the paradigm shift toward a learning experience that prepares students to be ready and willing participants in the global economy of tomorrow is fully in swing in schools everywhere. It remains though, that even with these strides toward progress, vast numbers of students still graduate without the skills necessary to fill an ever-increasing job pool of STEM-related careers.

In an attempt to better understand student achievement in mathematics, this literature review will focus on the ways in which students are evaluated in the mathematics classroom. Student assessment methodologies and grading paradigms are a popular topic of exploration as those in the field of education seek to find ways to optimize academics and improve student
achievement. Relevant research has shown that in addition to the way students are taught, the way in which students are evaluated for mastery can have an impact on their perceived and actual achievement.

Before examining the way in which students are graded, it is important to gain a better view of students’ mathematics achievement as a whole. This review will examine longitudinal trends throughout the years as the trends relate to student assessment, measures of student learning, and grading schema as a function of overall student achievement in the mathematics classroom. In addition to this, the review will compare student assessment metrics across different groupings (e.g. state and country) to establish relative performance baselines. Before seeking ways to improve the current situation, it is important to paint an accurate picture of the current state of affairs. It will then focus upon the inherent problems presented by a society lacking in its mathematics skills, and what other countries are doing to help remedy the situation.

The review will then turn its focus to traditional teacher-assigned grading systems, those based on percentages and overall letter grades. These are the standard de-facto instruments of measure and are widely prevalent in mathematics classrooms today as they were also years ago. This review will seek to clarify the way in which this system is used, the components and attitudes that comprise and influence traditional grades, as well as to identify its benefits and consequences. It will also explore the popular perceptions surrounding percentage-based systems, and their influence on the system’s continued use and popularity in schools throughout the United States. Traditional omnibus grading has been used in classrooms for close to a century and it the factors contributing to its ongoing perpetuation could shed light upon its influence on student achievement.
This review will then examine the increasing advent of standards-based grading systems as a reform-based change in the mathematics classroom and classrooms in general. Efforts to improve different aspects of the classroom – everything from the teachers to the desks – have been made in an attempt to solve the problem of dwindling student achievement. Non-conventional approaches to teaching and learning are increasingly gaining support as the tried and true methods of yesteryear are failing to appeal to the current generation of students. Grading students with rubrics on overall mastery of concepts is a relatively new practice, and is seen to align closer to the aims and methods of standards-based curriculum. The focus of this review will include the methods of its use, the motivation for implementation, and the benefits and consequences of its use by teachers and students.

The need for increased student achievement in mathematics is as apparent today as it has ever been. While much focus of the achievement debate has been upon the instructional and pedagogical methodologies employed by teachers in the classroom, a smaller subset of that conversation has been focused upon the way in which students are evaluated. Letter-grades and percentages serve as a uniform way to evaluate student’s learning, and seek to express a measure of mastery of the content area. As educators collectively work toward more authentic means of instruction and student engagement, similarly, that approach is being explored with respect to classroom evaluation paradigms. The measuring of student progress has and will always be a largely important component in how we determine student and teacher success. Keeping that in mind, it becomes important to be certain we are using the most authentic and representative metric possible to do so.

Mathematics teachers strive for their students to be highly literate in the content delivered to them in the classroom. Successful students are bred in successful classrooms that exhibit the
constructs of an environment that is conducive and supportive to student achievement. With that in mind, everything that takes place in the classroom should be closely examined when exploring avenues for increasing student achievement. Pedagogy is typically the component first examined – with good reason – but this review of the literature supports evidence that educators need to sometimes look beyond that. Policies and procedures dictating the ways students are evaluated in the classroom, in both a formative and summative means, can influence the ways students achieve, and perceive their achievement. These achievements then correlate directly with educational attainment.

The problem presented herein is that student achievement in the mathematics classroom suffers negatively from a traditional percentage-based grading schema. In turn, students are ill prepared for sequential post-secondary studies requiring mathematical proficiency. To gain a better understanding of the existing body of literature, this literature review will explore educational research regarding mathematics achievement, traditional grading schema, and standards-based grading schema.

**Student Performance in Mathematics**

The research is clear that United States students demonstrate a fundamental lack of basic mathematics skills upon entering high school (Silver, 2014) with a large majority of students failing to perform on grade level. The National Assessment of Educational Progress (NAEP) paints a picture of student performance in both reading and mathematics in fourth and eighth grade across the United States. According to data from the 2013 Nation’s Report Card, 64% of eighth grade students score less than proficient in mathematics (National Center for Educational Statistics, 2013). Students that are not proficient typically end up in some form of remedial math class that reinforces basic skills (Brantlinger, 2013) in an effort to better prepare them for future
math classes, and in turn give them skills needed to graduate and subsequently enter the workforce. Research suggests that upon graduation, a majority of students still lack basic skills with a majority requiring remedial assistance in their mathematics studies (Bahr, 2013) to be successful at a post-secondary level.

When examining the performance of United States students with respect to the rest of the world, students are also lacking in performance. Students in the United States score on average 75 – 100 points lower than students from countries with similar GDP (Lee and Huh, 2014). This deficiency in mathematics ability produces students that will have difficulty competing in the globally competitive marketplace (Silver, 2014) that exists today.

Students in North Carolina, when compared to their counterparts in the rest of the United States, are admittedly average performers (National Center for Educational Statistics, 2013). North Carolina ranks 26th in the nation on the eighth grade NAEP test, with 36% of students testing as proficient – an identical figure to that of the country as a whole. Viewed from any angle, a situation in which only a third of participants are successful warrants cause for some sort of drive for improvement.

The importance of mathematic literacy with regard to post-secondary success can not be overstated. Multiple research studies show that students that are proficient at mathematics and in turn complete a more rigorous course of study are much more likely to have substantially higher earning potential (Rose and Betts, 2001) upon exiting high school. Science, Technology, Engineering and Mathematics (STEM) based careers are growing at four times the rate of total employment over the past twenty years (Hossain and Robinson, 2012) and qualified individuals are needed to fill those roles. With a deficit in US graduates, these high paying stem careers are
typically filled by foreign-born workers (Hossain and Robinson, 2012) that have learned the requisite skills at foreign educational institutions.

In Germany, educational initiatives are being put into place to help German schools and universities produce more STEM-ready graduates. A trend towards integrated science education, rather than separate disciplines, can be seen beginning to gain momentum in elementary schools (Eilks and Markic, 2013). “In science as well as in mathematics education, standards and governmental syllabi suggest a greater orientation towards everyday life contexts and socio-scientific issues as the starters for science learning” (p. 1). Rooting STEM instruction in tangible relevant contexts increases student participation, retention, and builds enthusiasm more than abstraction and rote learning.

India, one of the world’s largest suppliers of STEM workers, started very early with its push toward training qualified graduates in mathematics and science. As early as the 1950s, science and mathematics literacy was seen as vitally important to the economic and social prosperity of Indian culture. (Kingdon, 2007). Incorporating the mastery of math and science into students’ first years of primary schooling is largely reflected in the India’s national standards (Government of India, 2002). Also reflected is the need for student-centered learning, with teachers playing the roles of facilitators as students construct their own knowledge through inquiry and experience.

When looking at increasing secondary mathematics achievements, reform efforts are characteristically based upon advances in pedagogical and learning theory (Marzano, Gaddy, and Dean, 2000). This is consistent with research showing that the quality of instruction from a teacher has a large impact on the achievement of students (Jacob, 2012). While this argument is
commonplace, increasing research is being made with regard to the way in which students are evaluated in the classroom.

There is a growing body of research that suggests a clear benefit to the utilization of meaningful and authentic assessments in the classroom (Cox, 2011). Although many definitions of authentic assessments are given in the literature, all include components of collaboration, application, and task-based learning (Cumming and Mawell, 1999; Lund 1997). Similarly, assessments that incorporate 21st century skills and focus on a student-centric accountability model of learning show correlation with increased test scores, increased student engagement, and higher levels of student motivation (Antonenko, Jahanzad, and Greenwood, 2014).

**Traditional Grading Systems**

Traditional grading systems, as they exist today, model a system adopted by Yale University in the late 1700s (Biggs, 2008) that separated students into 4 ranked groups based on their achievement level. Yale later adopted a numerical scale in lieu of group labels, with other colleges and schools subsequently following rank. Controversy over these shifts in thinking about student evaluation was rampant (Guskey, Swan and Jung 2001) and still continues to this day. It has become common today to question the role letter grades play in a student’s educational journey. Amongst educational researchers, there is some consensus that grading in its current form robs students of vital components at the core of education: passion, complexity, and depth of learning (Wisehart, 2004).

Although they are prevalent at every level of schooling, research has shown that traditional grading systems (teacher-assigned grades that are based upon an overall percentage and letter grade) tend to not accurately reflect what students know (Graeber, 2006). This is attributed to several different factors at play. First, task based applications as such do not lend
themselves easily to grading conventionally using an overall percentage and letter grade. Teachers are in a constant state of adjustment with regard to their instruction, revising and adapting to what students know (Gearheart and Saxe, 2004). This can make measurements of learning difficult when a percentage or letter grade is the only metric in place, and it does not accurately convey knowledge or mastery (Cox, 2011). Moreover, inconsistencies from teacher to teacher with regard to difficulty, class policies and procedures, and their own belief structure factor into the way students are assigned a percentage and letter grade (Cox, 2011). Research has shown that teachers often employ systems of grading and evaluation with which they are most familiar, often times mirroring educational experiences the teacher previously had (Cox, 2011).

Similarly, traditional grading systems often factor in non-content standards such as timeliness, neatness, and attendance into a grade leading to a distortion in seeing what has actually been mastered (Deddeh, Main and Fulkerson, 2010). Often referred to as “hodge-podge” grading practices (Cross and Frary, 1999), letter grades assigned to students by teachers are often distorted and diluted by these factors, leading to inflated grades and grades that often do not reflect what the student actually knows. For these reasons, Shephard, Owen, Fitch and Marshall (2006) argue out of the multitude of metrics available to schools to measure student growth and achievement, traditional teacher-assigned grades are not a reliable one, limiting their scope of usefulness to parents, employers, and college admissions counselors. Letter grades assigned by teachers have a high-face value in the eyes of many, regardless of the components that contributed to the creation of the grade itself.

Grades exist to paint a picture of the level of achievement attained by a student (Allen, 2005). Complexities in the assignment of grades and grading schema have led teachers to adopt a one-size fits all approach that typically mirrors what the teachers themselves experience in
school as a student (Guskey, Swan and Jung, 2011). Coupled with the facts that teachers at the same school come from varied educational backgrounds, and that historically, the majority of experiences do not align with current research and pedagogical advances (Guskey, 2011). These factors combine to drive the body of research to agree traditional teacher-assigned grades are extremely poor predictors of academic mastery (Evans, 1976; Shepard 2006; McMillan, 2003).

The consensus on the lack of utility of traditional grading systems has led educational researchers to implore teachers to strive to create more authentic assessments that focus more upon standards (Carr, 2000).

The aforementioned lack of utility has been demonstrated by researchers through comparison of traditional teacher-assigned grades and standardized achievement metrics. Brennan, Wenz-Gross and Siperstein (2001) found a significant gap existed between the grades that students earned in class from their teachers and the grade that student’s earned from a standardized assessment. The argue further that this is a direct product of grading practices in the classroom that are inconsistent and “fuzzy” in nature (2001). Busick (2000) argues for teachers to strive to close the gap as a service to everyone who so heavily relies on traditional teacher-assigned grades, holding them in high esteem. Regardless, much resistance is exhibited as teacher surveys indicate an overwhelming preference toward traditional grading (Shepard, 2006) as teachers seek to find a way to hold students accountable for not only their knowledge, but their adherence to educational conventions as well.

The disconnect that is evident between traditional teacher-assigned grades and their standardized metric counterparts lends itself well to the existing body of research that attempts to diminish the value of traditional classroom grading. Lekholm and Cliffordson (2008) point out that while these two different metrics remain divested of one another in the United States, other
countries have taken initiatives to help close that gap by using standardized test data to troubleshoot traditional teacher-assigned grades (Lekholm and Cliffordson, 2008).

Another thing that contributes to the apparent dilution of traditional teacher-assigned grades is the subjective nature of grades themselves. Opportunities for bias introduce themselves when teachers assign grades that include things not easily measured such as neatness, adherence to project guidelines, and creativity (Jae and Cowling, 2009). Weimer (1998) admits on behalf of teachers everywhere, that although “all of us aspire to treat students fairly and equally… Unfortunately, despite our aims of objectivity and fairness, our humanity sometimes gets in the way. It is only natural to ‘like’ some students more than others” (p. 5). Teachers, much like any other subset of humanity, are bound to exhibit varying levels of bias with respect to gender, ethnicity, and perceived level of ability (Jae and Cowling, 2009).

It follows logically that grades shown to be rooted in subjectivity, often diluted by non-academic factors and clouded by inherent subconscious bias, do not serve as an accurate measurement of what students know (Evans, 1976). The body of literature suggests that a focus upon improving grading as a whole should make an attempt to eliminate or limit these factors that contribute negatively to the face value of a traditional teacher-assigned grade (Darling-Hammond, 1994). Teachers uniformly express concern about grading students fairly (Rex 2005; Yip and Cheung 2005) but due to their backgrounds and experiences, they often offer differing opinions on what fair grading actually is (Zoeckler, 2005).

**Standards-based Grading Systems and Reform**

The usefulness of grading is very much dependent upon students, teachers, and parents all maintaining a shared perspective of what the grade is attempting to communicate. Austin and McCann (1992) viewed grades as a language that all stakeholders must speak fluently to
effectively interact with one another, while Lindblom and Cohen (1997) emphasize the importance of the shared experiences and understandings of the group when sharing student progress through the use of grades. This can create a problem in that while teachers and parents undoubtedly have both experienced backgrounds in which similar grading systems have been used to chart their efforts, there can be much misunderstanding and misalignment in the connections and conclusions that teachers and parents ultimately draw (Airasian, 1998). The literature shows that reforms in grading have focused on this disconnect.

Efforts to make assessments and grades more representative, transparent, and authentic are driving the conversation behind grading reform. Authentic student assessments serve as a means for students to collect and use information to make decisions and solve problems as well as serving to reliably inform teachers the abilities of their students. (Kulm, 2013). With the advent of the Common Core Curriculum Standards for Mathematics, emphasis on problem solving and synthesis of knowledge has increased significantly. Curriculum focusing on higher levels of rigor should have assessments structured to do the same. More focus is being placed on schools transitioning from tests and quizzes, and toward learning tasks and collaborative projects (Graue, 2005). These newly explored visions of assessments also help to address the ambiguity created by the juxtaposition of competence and proficiency (Silver and Keaney, 1995).

Additionally, authentic assessments should help teachers gain information and insight about the process of teaching and learning (Darling-Hammond, 1994) by creating a constant cycle of reflection and revision. Context can also play a large role in making assessment more authentic and relevant to students (Lajoie, 1995). By reflecting the way learning and application often happen outside the classroom – namely group-based and grounded in a real-world frame of reference – assessment becomes more representative of learning and mastery (Lajoie, 1995).
Changes in the mathematics classroom with respect to the way lessons are designed are becoming more aligned to the recommendations of the National Council of Teachers of Mathematics by placing an emphasis on 21st century skills including problem solving, decision making, and collaboration (Bay, Reys and Reys, 1999). In the same classrooms that are leading the way pedagogically, assessment reform is also taking place (Iamarino, 2014). Much like the content itself, standards-based grading places an emphasis on mastery of the individual standards in mathematics (Guskey et al., 2011). By eliminating items that are not grounded in the mathematical standards, a more accurate representation of student achievement is gleaned (Deddeh et al., 2010). Standardizing the content in the mathematics classroom has also led to an increased adoption of common assessments (Cox, 2011) aligned to content standards. These assessments aim to decrease the variation in measurements of student achievement from teacher to teacher, and from teacher to student performance on standardized tests.

Guskey and Jung’s (2006) research identified six key areas in which grading could be reformed to help it be more meaningful and representative of student knowledge: making a conscious attempt to assess most objectives in the state standards; focusing grading on student achievement, removing effort from their appraisals; creating assessments to focus on state standards; identifying the objective or strand associated with each assessment item to track student performance skill-by-skill; being concerned with attainment of state standards rather than the skills outlined in the text; and, using only assessments administered toward the end of a unit, when students had developed proficiency, to assign grades. Shippy, Washer and Perrin (2013) reach similar conclusions and make an argument for an assessment plan that is modeled after curricular plans, one that is based upon standards.
Standards-based grading systems serve to measure student progress toward their achievement of clearly outlined course objectives (Heywood, 2014). Although it is aligned to course standards, researchers argue that standards-based grading remains very personalized in that it measures students using a criterion-based approach (against the standards themselves) rather than a norm-based (against the performance of other students) approach (Carberry, Siniawski, and Dionisio, 2012). This type of measurement of achievement has been viewed favorably by students as more transparent and clear in making students aware of expectations (Schmoker and Marzano, 1999) while eliminating relative comparison with other (Sadler, 2005).

This increase in clarity for students and for teachers brings with it a clearer picture of the student’s strengths and weaknesses (O’Connor, 2007) and allows teachers, parents, and students to better focus on what needs to be done to ensure success in the classroom. The consistency of standards across subjects and school years helps create better longitudinal measurements over time that are interdisciplinary and can be used to reinforce efforts for remediation and enrichment (Tognolini and Stanley, 2007).

The interdisciplinary nature of a standards-based approach to assessment also allows for increased opportunity for teacher collaboration (O’Connor, 2007). Schmoker and Marzano (1999) argue that the overlap seen in standards between curriculums will allow teachers to plan for more comprehensive educational experiences that allows students to demonstrate their mastery in multiple disciplines. Similarly, these experiences lend themselves well to real-world applications that when incorporated into the classroom, create an enriched learning environment for students to develop critical thinking skills (Doherty & Hilber, 2008).

One of the most common ways standards-based grading is incorporated into classrooms is through the use of a detailed rubric that communicates goal level mastery and assigns students
a rating on a sliding scale. Rubrics are often used as a focused method of assessment for learning that aim to inform all stakeholders (Wollenschlager, Hattie, Machts, Moller, & Harms, 2015). The growing popularity of rubrics as an evaluation tool can be attributed in part to the ease of application to a wide-range of tasks and skills (Thompson & Senk, 1998) and the ability of rubrics to provide rich feedback that pays attention to learning objectives and the achievement of those objectives (Stevens & Levi, 2004).

Critics of a standards-based approach to grading argue that it detracts from teacher individuality (Cox, 2011) and can create an environment where students are not held accountable for homework assignments (Bursuck, Polloway, Plante, and Epstein, 1996), attendance, or other non-achievement measures (Allen, 2005) that may exhibit a large influence on the overall academic performance of students. Stiggins (2001) and Wiggins (1996) propose a grading system in which students earn separate marks for content and study habits, arguing that such a system would give parents more detailed feedback on how their students could improve skills that would directly translate to an increase in content mastery. They also argue that this is easier upon teachers, because the act of consolidating a large number of diverse factors into a single letter grade is often tedious and cumbersome.

Another factors to consider when reforming the traditional grading system include a sizeable commitment of resources to the education and training of teachers (Allen, 2005), students, and parents in its interpretation. Letter grades are simple, familiar, and easy to read and compare. Although standards-based reports would provide more detail, there’s also more effort in deciphering them to a useful end result. Critics argue that a failure to reach a consensus on the calculation of an overall score on standards and strands, or “true score” to which Marzano (2006) refers, introduces the same levels of ambiguity seen with teacher-assigned traditional grades.
Advocacy Argument

The data identified by the body of literature indicates a downward trend in student achievement in mathematics in the United States. Students exhibited low levels of proficiency when compared with their counterparts throughout the world. While the research reviewed did not address the direct cause for students’ deficit of ability, it does present the case that logical interventions will be necessary to address the issue. While most reforms have been pedagogical in nature, the research reviewed argues the case for expanding that scope to student evaluation as well. While a standards-based approach to curriculum has been effective in increasing student achievement in research samples, it must also be shown that those results can be translated to the whole population. At the same time, a body of evidence must be created that establishes standards-based grading in a parallel fashion.

A large body of literature exists on student evaluation, assessment, and grading paradigms in the secondary school classroom. It is clear from the research that while authentic forms of assessment have been shown in many cases to provide a more meaningful measure of student learning, beneficial to students and teachers alike, much resistance has been met from all stakeholders involved in transition to what is viewed as radically new way of thinking about grades. While the roots of this resistance have been explored, there is little empirical research that has been conducted with regard to quelling this resistance and easing the transition of all parties involved in the process. If an understanding has been reached regarding the root cause of a problem and its sources have been identified, then logically it would follow that future research should be devoted to remedying that problem.

It should also be noted that in a large number of studies focusing upon grading system reform, similar reforms were also taking place with regard to instruction and teaching. It is not
clear if the benefits of the grading systems were a direct result of the shift in grading or the shift in pedagogy, as it seems that most studies enacted the two reforms simultaneously. In doing so, it becomes difficult for researchers to discern the benefits from one reform or the other. This places great limitations on a sizeable portion of the available literature, creating a need for further research with a more narrowed focus.

Furthermore, while researchers have reached a general consensus on the benefits of standards-based grading and the advent of more authentic forms of assessment of students, these broad generalizations based on observation in schools in general may not translate specifically to the mathematics classroom. The majority of studies reviewed herein enacted the reforms on a school-wide level and did not stratify their results by subject area. It could be argued that it is not clearly known the effects such a grading system has on a mathematics classroom. Mathematics in itself is entirely objective in its nature, and may not lend itself well to a standards-based grading system that implements levels of subjectivity in its rubrics. Therefore, it is important that future research focus upon the measurable and/or observable implications that such a system would present in a secondary mathematics classroom, as very few bodies of literature exist upon the subject.

**Conclusion**

It is clear from the existing body of research that there is much debate surrounding the idea of student achievement as a factor of several different variables. A large emphasis has traditionally been placed upon the reform of instructional methodology in the classroom, and increasingly, a growing body of research is exploring reform of evaluative practices used in measuring students’ achievement.
Percentage and omnibus-letter grading are still employed in the majority of schools in the United States as a product of teacher familiarity and ease of implementation. However, it can be argued that appeal to tradition is not always the best approach, and that traditional grading systems do not best serve the students’ and teachers’ needs. The fuzziness introduced into the grading process through subjectivity, bias, and ambiguity creates a grade that while convenient and familiar, does not create an accurate portrayal of student mastery. A grading system that more closely mirrors the reform-based measures taking place in mathematics pedagogy and student learning may be more conducive to increasing overall student achievement. Traditional grading systems appear to suffer in this area and sometime create the perception that grades do not truly reflect what a student has mastered. This creates a disconnect that makes it difficult for everyone involved, and devalues the authenticity of grades. This problem is exacerbated by the various life-impacting institutions that still use grades and grade point averages as part of their evaluation criteria. Ease of comparison and ability to quickly reduce prospects to a ranking undoubtedly contribute to the perpetuation of this system.

Much like curriculum has increasingly become standards-driven in an attempt to drive instruction in a more concrete direction, it would logically follow that grading should move along a similar route. Reforms in the classroom that aid students in reaching their educational goals are much more difficult to quantify in results if the instruments with which we measure are inaccurate or misleading. Additional research is needed to see if a difference in grading system alone exhibits a significant impact on student achievement. The goal of this study is to discern the perceived and actual benefits from replacing the traditional grading system with one that is more aligned to a standards-based classroom approach.
Chapter 3: Methodology

**Research Questions**

To what extent does the implementation of a standards-based grading system affect the overall student achievement in a statistically significant way in an Algebra 1 class at a rural high school in North Carolina? Similarly, to what extent do students gain motivation or academic purpose when given more specific feedback on their performance that is standards-based?

**Purpose of the Questions**

Bandura's (1977) social cognitive theory emphasizes the importance of relationships and feedback in determining student's motivation and self-efficacy of their mathematic abilities. These interactions and shared experiences are experienced by the teacher and student in different ways in both traditional grading systems, and standards-based grading systems. By examining the achievement of two groups of students with similar abilities, the researcher hopes to establish a statistically significant impact, if it exists, of grading systems on student post-test achievement levels. The literature is in agreement that traditional grading systems often do not reflect student mastery (Graeber, 2006), contain a variety of non-content components (Cross and Frary, 1999), and vary wildly with teacher inconsistencies (Cox, 2011). This leads the researcher to believe that student motivation and self-perception would be affected in a negative way using a traditional grading system. Standards-based grading eliminates these issues in lieu of a rubric-like format (Guskey and Jung, 2006).

**Discussion of the Variables**

Independent Nominal Categorical Variable: Classroom Grading System: this variable would describe the grading system employed in the Algebra 1 classroom - either Traditional Letter Grading, or Standards-based grading for this study.
Co-variates Interval Variable: Student Pre-Test Achievement: This variable would measure student achievement levels as measured by an Algebra 1 pre-test instrument.

Dependent Interval Variable: Student Post-Test Achievement: This variable would measure student achievement levels as measured by an Algebra 1 post-test instrument.

Hypothesis / Rationale

Algebra 1 students that are assessed using a standards-based grading system will experience higher levels of growth as measured by post-test instruments that similar counterparts in a traditional grading system. The literature reviewing standards-based grading creates an opportunity for students to receive more authentic and descriptive grades, form better perceptions of their strengths and weaknesses, and subsequently experience higher levels of achievement. Serving as a feedback and scaffolding device throughout the semester, standards-based grading provides a means for students and teachers to be more reflective and critical of teaching and learning.

Research Design

Since true randomization of research participants is nearly impossible in a school setting, the researcher has chosen a quasi-experimental research design. Qusai-experimental design is often used in education and social science settings when true randomization of subjects is easily obtained or feasible (Muijs, 2004). Four different classes of Math 1 students at the same high school, similar in composition with regard to educational attainment, socio-economic status, and mathematical abilities. Two of the classes will be the experimental group, and will be evaluated throughout the semester using standards-based grading. The two other groups will be the comparison group, and be evaluated throughout the semester using familiar traditional grading systems utilized throughout the rest of the school. A quasi-experimental design, one in
which measurements and observations are made before and after the treatment, will be utilized to help identify the impact of grading upon the achievement of the student participants (Fraenkel, Wallen and Hyun, 2012).

**Justification for the Design**

Quasi-experimental studies, unlike true experimental studies, do a better job of simulating real-world classroom environments (Muijs, 2004). The research environment in which the study will be conducted allows the comparison and experimental groups to be very similar in their make up, greatly reducing the chance for extraneous interference. It will be important to gather as much information as possible about the members in each group to control for any anomalies that may otherwise affect the outcomes. Conducting the research in a classroom setting helps to generalize the results in the event that correlations are found. Since they experiment was conducted in a classroom, the results would more likely translate to the classroom than if conducted in a laboratory setting.

**Implications**

Several implications exist for this research study. Should the null hypothesis be disproved by demonstrating statistical significance of standards-based grading on student achievement, this would demonstrate the need for further study to increase confidence of generalizability, and hopefully encourage more teachers to adopt a standards-based grading approach in their mathematics classroom. Should the null hypothesis be confirmed showing that there exists no statistically significant relationship between standards-based grading and student achievement, this may spark ideas for refinement, revision, and other grading reform efforts to be implemented and researched in a similar effort.
Limitations for the Design

The limitations of a quasi-experimental design are found in the effort to keep the experimental and comparison groups as close to each other in all aspect as possible. A lot of effort will need to go into ensuring educational, socioeconomic, and demographic measurements are reasonably comparable between groups. Not doing so creates variables unaware to the researcher that can threaten the validity of the study and cause concern when trying to infer conclusions from the data.

Extraneous variables associated with student achievement could possibly have effects on the results of the study – including socioeconomic status, mathematical aptitude, family life, and social factors.

Threats to internal validity such as unanticipated events happening during the study, changes in the developmental levels of students during the experiment, groups of low or high students regressing toward the mean, participants dropping out of the experiment or transferring classes, rival classes competing against each other in an effort to “one-up” the other, or the Hawthorne effect which can alter student results in unexpected ways just because they know they’re participating in research.

Site and Participants

The research site is West Lincoln High School, a secondary public school in rural North Carolina located approximately 40 miles from Charlotte. The student population consist of approximately 950 students, 80% of which are White, 8% of which are Hispanic, and 8% of which are African-American. West Lincoln High School’s rolling three year average for proficiency (Level 3 score) in Math 1 is 75%, with a college and career ready proficiency (Level 4 score or higher) of 61%. The site was chosen due to familiarity to the researcher, and the
homogeneity of its student population with regard to socioeconomics and demographics. -
Additionally, West Lincoln High School also sends a large portion of its incoming freshman
class through Math 1 (85 - 90% of students) providing a large population from which to choose a
sample. -

Unlike other math courses at the secondary level, Math 1 is a universally required course
for all students and is not considered to be theoretical or an upper-level course that could introduce struggles for some students who have yet to become formally operational in their thinking. Students in this course are engaged in critical thinking, problem solving, and cooperative learning, using basic mathematical skills as their methods.

The research participants in the study will be four classes of Math 1 classes during the Spring semester at West Lincoln High School, comprised of approximately 25 students each, out of a total of 125 students enrolled in Math 1 during the Fall semester. The target population are students who have completed Math 1 since 2012, the year that the new Common Core mathematics curriculum was implemented. The research question guiding this study seeks to better understand the effects of grading schema on Math 1 student achievement, and by choosing students that are actively enrolled from a population with similar demographics and curricular goals, the analysis of pre- and post-test scores will allow the researcher to determine whether measurable effects exist.

The criteria for creating the sample is as follows: 9th grade student at West Lincoln High School, enrolled in Math 1 during the Spring 2018 semester, Although this is a quasi-experimental study utilizing an experimental group and a control group, the role of the classes participating in the study will be randomly drawn from a hat.
The role of the researcher in this study is limited, as classes participating in the study will not have direct contact with the researcher. The role of the researcher in establishing the operating parameters for classrooms in the experimental and control groups will be seen in the daily operation of the classroom. The researcher, being an expert in the treatment being applied, will be directly responsible for instructing teachers that are carrying out the treatment upon the experimental group.

**Protection of Human Subjects**

Students participating in this study are protected from undue harm or negligence in the classroom by ensuring the curricular goals and educational benchmarks set forth by the North Carolina Standard Course of Study are present in both the experimental and control groups, regardless of grading schema. Traditional grading systems, based omnibus letter grades and percentages, are the de facto metric used when evaluating student mastery in the classroom. Continuing to do so in the control classroom places no risk to that group. Student and teacher names will be redacted and replaced with pseudonyms to protect their identities.

While the impact of standards-based grading in the experimental classroom has yet to be empirically quantified, students in this group are still receiving the same instruction, held to the same instructional standards, and engaged in the same higher-order thinking tasks that Common Core brings along with it. Using a new metric to track student mastery does not place these students at any risk of not receiving the information they need to be successful in the classroom. The researcher's positionality as a mathematics teacher at West Lincoln High School, as well as prior experience with standards-based grading in lieu of traditional grading systems and those subsequent opinions and biases formed therein, underscore the importance of the researcher having little to no role in the instruction of the experimental and control groups in this study. For
that reason, the classes participating in the study will receive instruction from another Math 1 teacher at West Lincoln High School.

Instruments

The data that will be collected from pre- and post-test scores is ratio data, because there is an absolute zero value (0% correct) and differences between values can be compared meaningfully. The difference between the pre-test and post-test allows the researcher to establish growth metrics for students in the control and experimental groups.

A pre-test will be given at the beginning of the academic semester to all students enrolled in the control and experimental classes to establish a baseline level of aptitude for the course curriculum. This test would be an instrument measuring mathematical ability with regard to eight grade math skills. At the end of the semester, students will be administered the an End-of-Course test and growth scores will be computed by calculating each student's difference in scores. The pre-test score will be used as a co-variate measurement in the ANCOVA for this study. ANCOVAs are suitable for experimental studies to control for factors that cannot be randomized, but can be measured on an interval scale. Additionally, the use of a co-variate pre-test baseline as a control variables dictates the need for an ANCOVA. An F-Test will show if the co-variates exhibit any significance upon the final post-test scores.

The pre-test scores used to establish prior academic achievement mastery levels will be gathered through the use of a psychometrically balanced test previously used in North Carolina as an End-of-Grade examination for eighth grade, but has since been released for teachers to use in preparation for exams. These exams are public domain, freely reproducible, and available for use by teachers across the state of North Carolina.
The post-test score will be established using the North Carolina End-of-Course examinations. These tests are developed by teachers in North Carolina, and aligned to the NC Math 1 standard course of study and pacing guides utilized throughout the state. Test items normed through a field testing process in which items are assigned a level of difficulty, known as a p-level. Tests are then balanced at the item level to ensure all form versions are of equal difficulty. All tests undergo a thorough review process comprised of 22-steps in which there are checks for bias, readability, and best-fit of math standards to each item.

The focus group interviews serve to clarify the results from the quantitative instruments and better answer the research questions set forth in this study. The responses from interviewees will be coded by the following themes: achievement (grades, quality of work, understanding, comprehension), communication (teacher feedback, student reception, clarity), and student improvement (growth, remediation, and attitude).

Content validity is ensured with alignment to Math 1 standards for North Carolina. Criterion validity is ensured through the field testing, norming, and balancing of each form to align student mastery with test outcomes. Construct validity is ensured through a multi-dimensional picture of student growth, comparing beginning and end of semester scores, to get a better picture of how students have growing in each of the classrooms. Internal consistency reliability will be apparent by the emergence of similar results within the control group and within the experimental groups. Various statistical tests measuring internal consistency may be of use here to affirm reliability.

Reliability ratings for the North Carolina End of Course tests for Math 1 are as follows: Form A, 0.91, Form B, 0.91, Form M, 0.90, and Form N, 0.90.
**Procedures**

Four classes will be selected randomly to participate in the study (two as control, and two as experimental), the classes will be given a pre-test during the first week of school to establish a baseline for mastery. Student names will be masked, and scores will be sorted by randomly generated ID numbers. Students will participate in identically designed curriculums throughout the semester, with the only difference being the way in which they are evaluated and given feedback through differing grading paradigms.

In control classrooms utilizing traditional grading, students will be assessed using an overall proficiency ratio of points earned to points possible. This will then be converted into a percentage which will then be used to communicate the percent of the content mastered to the students. Assessments for each unit will be returned with markings indicating which problems were incorrect. The overall percentage will be assigned a letter grade in accordance with West Lincoln High School’s ten-point grading scale, which in turn will be recorded on the top of the test. These grades will be recorded in the class gradebook and will be communicated via interim and grading period ending reports.

In experimental classrooms utilizing a form of standards-based grading, the researcher and participating teachers will develop standards-based rubrics that will reflect the learning goals for each unit in the class. The rubrics designed for each unit will be embedded within the assessments, with exercises grouped into similar categories to aid in ease of assessment. Students will be assigned a rubric rating in each learning goal ranging from 0 (not attempted) to 4 (mastery). Additionally, students will also receive a performance summary sheet with each examination that summarizes the learning goals from that assessment and outlines their performance rating in each. To aid in the comprehension of this new grading system, students
and parents received an informational handout at the beginning of the semester detailing grading methodology and answering frequently asked questions.

During the last week of the semester, students in each class will be administered the same pre-test from the beginning of the semester, with scores being sorted using the previously assigned ID numbers. Value-added measurements will be calculated using the difference between the pre- and post-test scores.

To ensure the validity of the survey instruments (i.e. pre- and post-tests) forms will be retyped, relabeled, and have the origin masked. Item orders and answer choices will be shuffled. Tests will be collected, and answers to test items will not be made available until after the post-test has been graded.

Different teachers will participate in the study’s implementation, increasing confidence in implementation validity. This helps minimize selection bias by giving each group an identical teaching experience. The control group helps minimize the threats to internal validity by creating a situation where history or maturation effects could affect the results. Using the entire 9th grade as a population from which the sample is taken lessens the likelihood of statistical regression, as no group will be wholly high or low performers that would regress toward the mean.

At the conclusion of the research study, students and teachers in the experimental groups will be assembled into a focus group that aims at gathering feedback regarding the usefulness of the feedback given to them regarding their performance throughout the semester. The aim of this qualitative component is to help further clarify and understand the results of the quantitative study. Students and teachers will provide insight as to the standards-based feedback’s effect on motivation, persistence, and practicality as part of their educational experience. These interviews will be transcribed, coded, and analyzed for patterns and shared connections between responses.
Validity, Reliability, Generalizability

The North Carolina End of Course exam for North Carolina Math 1 is a standardized test given to all students enrolled in NC Math 1. Practice tests are made available to teachers through released forms, that were previously used for evaluation of students. The achievement tests used for this course are aligned to the NC Common Core math goals that guide the pacing guide and standard course of study for NC Math 1. This ensure content validity. Since these tests are field tested on students to create a baseline for levels of achievement, this ensures predictive criterion validity. Certain achievement levels demonstrate a students knowledge in NC Math 1.

North Carolina End of Course exams undergo a 22-step review process to eliminate bias, enhance clarity, and ensure alignment to content standards. This process helps to ensure reliability by lessening the chance of random error. The review process ensures inter-rater reliability, while field testing allows for test-retesting reliability.

Generalizability can be described as making predictions based on past observations. In this study, the sample of students is representative of the demographics of North Carolina, and students currently enrolled in NC Math 1 throughout the state. West Lincoln High School is a rural school located 45 miles from Charlotte, North Carolina. To increase the confidence of generalizability, this study would need to be repeated at schools throughout North Carolina, with different educators responsible for the implementation. Larger sample sizes would lead to more confidence in the significance of the experimental treatment.
Chapter 4: Analysis of Findings

**Introduction**

The purpose of this research study was to determine the impact of grading methods on student achievement in secondary mathematics, as well as student perception of those grading methods upon their achievement. In doing so, the researcher gathered information from two separate instruments: a quantitative analysis of pre- and post-test scores in North Carolina Math 1 classes, and qualitative analysis of coded focus group interviews of both teachers and students participating in the research study. Taking both of these into account, the researcher aims to answer the following questions:

1. Does the way in which students are graded exhibit a measurable influence upon overall student achievement in a secondary mathematics class?

2. To what extent do students perceive grading systems as having influence on their achievement in secondary mathematics?

**Quantitative Results**

**Pre-test and Post-test Scores**

Pre-test scores were gathered from all participating classes at the beginning of the semester by administering an 8th Grade Math EOG released test. All students enrolled in Math 1 had taken a similar exam at the conclusion of the previous school year. The pre-test was administered before any instruction in Math 1 had begun. 72 students’ test scores were analyzed for this research study, with 39 students residing in the control classrooms, and 33 residing in the experimental classrooms. The student pre-test consisted of 38 exercises aligned with 8th Grade Math goals including the number system, expressions and equations, functions, geometry, and
statistics and probability. A calculator was provided. Results from the pretest can be seen in tables 1 and 2.

Table 1

**Student Pre-Test Results from Control Classrooms**

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>n=39</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scored 0-10% Correct</td>
<td>4</td>
<td>0.102</td>
</tr>
<tr>
<td>Scored 11-20% Correct</td>
<td>11</td>
<td>0.282</td>
</tr>
<tr>
<td>Scored 21-30% Correct</td>
<td>22</td>
<td>0.564</td>
</tr>
<tr>
<td>Scored 31-40% Correct</td>
<td>2</td>
<td>0.051</td>
</tr>
</tbody>
</table>

Table 2

**Student Pre-Test Results from Experimental Classrooms**

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>n=33</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scored 0-10% Correct</td>
<td>2</td>
<td>0.061</td>
</tr>
<tr>
<td>Scored 11-20% Correct</td>
<td>8</td>
<td>0.242</td>
</tr>
<tr>
<td>Scored 21-30% Correct</td>
<td>19</td>
<td>0.575</td>
</tr>
<tr>
<td>Scored 31-40% Correct</td>
<td>4</td>
<td>0.121</td>
</tr>
</tbody>
</table>

Before analyzing posttest scores using ANCOVA, it is important to check the strength of the co-variate. To do this, we will corelate it with both pre-test and group.

Table 3

**Correlation of Covariate**

<table>
<thead>
<tr>
<th></th>
<th>Pretest</th>
<th>Postest</th>
<th>Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pretest</td>
<td>Pearson Correlation</td>
<td>.335**</td>
<td>.072</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.004</td>
<td>.546</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>72</td>
<td>72</td>
</tr>
<tr>
<td>Postest</td>
<td>Pearson Correlation</td>
<td>.335**</td>
<td>.437**</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.004</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>72</td>
<td>72</td>
</tr>
<tr>
<td>Group</td>
<td>Pearson Correlation</td>
<td>.072</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.546</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>72</td>
<td>72</td>
</tr>
</tbody>
</table>

**. Correlation is significant at the 0.01 level (2-tailed).**
From our analysis, pretest is a good covariate shown by its correlation with the dependent variable postest, but not the independent variable group.

As evidenced by the one-way ANOVA conducted in the table above, there is no statistically significant effect from membership in either group, \( F(1,72) = 0.37, p = 0.55 \).

At the conclusion of the semester, students were administered an End-of-Course test aligned with North Carolina Math 1 standards including number and quantity, algebra, functions, geometry, and statistics and probability. The post-test consisted of 51 graded exercises and the results can be seen as summarized in the tables 4 and 5.

Table 4

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>n=39</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scored 30-40% Correct</td>
<td>2</td>
<td>0.051</td>
</tr>
<tr>
<td>Scored 40-50% Correct</td>
<td>4</td>
<td>0.103</td>
</tr>
<tr>
<td>Scored 51-60% Correct</td>
<td>7</td>
<td>0.179</td>
</tr>
<tr>
<td>Scored 61-70% Correct</td>
<td>9</td>
<td>0.231</td>
</tr>
<tr>
<td>Scored 71-80% Correct</td>
<td>7</td>
<td>0.154</td>
</tr>
<tr>
<td>Scored 81-90% Correct</td>
<td>6</td>
<td>0.128</td>
</tr>
<tr>
<td>Scored 91-100% Correct</td>
<td>4</td>
<td>0.103</td>
</tr>
</tbody>
</table>

Table 5

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>n=33</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scored 30-40% Correct</td>
<td>0</td>
<td>0.000</td>
</tr>
<tr>
<td>Scored 40-50% Correct</td>
<td>0</td>
<td>0.000</td>
</tr>
<tr>
<td>Scored 51-60% Correct</td>
<td>2</td>
<td>0.060</td>
</tr>
<tr>
<td>Scored 61-70% Correct</td>
<td>5</td>
<td>0.152</td>
</tr>
<tr>
<td>Scored 71-80% Correct</td>
<td>10</td>
<td>0.257</td>
</tr>
<tr>
<td>Scored 81-90% Correct</td>
<td>11</td>
<td>0.282</td>
</tr>
<tr>
<td>Scored 91-100% Correct</td>
<td>5</td>
<td>0.152</td>
</tr>
</tbody>
</table>

To establish a value-added metric for students in both groups, a student growth score was calculated by subtracting the pre-test score from the post-test score. These scores are summarized in the tables 6 and 7 below.
Table 6

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>n=39</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Growth of 0-10%</td>
<td>3</td>
<td>0.077</td>
</tr>
<tr>
<td>Growth of 11-20%</td>
<td>2</td>
<td>0.051</td>
</tr>
<tr>
<td>Growth of 21-30%</td>
<td>5</td>
<td>0.128</td>
</tr>
<tr>
<td>Growth of 31-40%</td>
<td>9</td>
<td>0.231</td>
</tr>
<tr>
<td>Growth of 41-50%</td>
<td>11</td>
<td>0.282</td>
</tr>
<tr>
<td>Growth of 51-60%</td>
<td>7</td>
<td>0.179</td>
</tr>
<tr>
<td>Growth of 61-70%</td>
<td>2</td>
<td>0.051</td>
</tr>
</tbody>
</table>

Table 7

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>n=33</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Growth of 21-30%</td>
<td>2</td>
<td>0.061</td>
</tr>
<tr>
<td>Growth of 31-40%</td>
<td>2</td>
<td>0.061</td>
</tr>
<tr>
<td>Growth of 41-50%</td>
<td>12</td>
<td>0.364</td>
</tr>
<tr>
<td>Growth of 51-60%</td>
<td>10</td>
<td>0.333</td>
</tr>
<tr>
<td>Growth of 61-70%</td>
<td>5</td>
<td>0.152</td>
</tr>
<tr>
<td>Growth of 71-80%</td>
<td>2</td>
<td>0.061</td>
</tr>
</tbody>
</table>

When analyzing the growth metrics, the control classes as a whole experienced less growth on average \( M = 39.56\% \), \( SD = 15.15\% \) than the experimental classes \( M = 50.51\% , SD = 14.96\% \). Before proceeding with the ANCOVA, to verify the homogeneity of the regressions, we will test the interaction of group and pretest.
Table 8

**Group and Pretest Interaction**

Dependent Variable: Postest

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected Model</td>
<td>.481^a</td>
<td>3</td>
<td>.160</td>
<td>9.006</td>
<td>.000</td>
<td>.284</td>
</tr>
<tr>
<td>Intercept</td>
<td>2.621</td>
<td>1</td>
<td>2.621</td>
<td>147.143</td>
<td>.000</td>
<td>.684</td>
</tr>
<tr>
<td>Group</td>
<td>.047</td>
<td>1</td>
<td>.047</td>
<td>2.661</td>
<td>.107</td>
<td>.038</td>
</tr>
<tr>
<td>Pretest</td>
<td>.154</td>
<td>1</td>
<td>.154</td>
<td>8.632</td>
<td>.005</td>
<td>.113</td>
</tr>
<tr>
<td>Group * Pretest</td>
<td>.002</td>
<td>1</td>
<td>.002</td>
<td>1.06</td>
<td>.745</td>
<td>.002</td>
</tr>
<tr>
<td>Error</td>
<td>1.211</td>
<td>68</td>
<td>.018</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>38.201</td>
<td>72</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>1.692</td>
<td>71</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

^a. R Squared = .284 (Adjusted R Squared = .253)

The interaction term was found to be not significant \((F(1,71) = 0.106, p = 0.745)\), thus the assumption of homogeneity of slope was not invalid.

The subsequent ANCOVA test was then used to examine the significance of group membership controlling for pretest as shown below in Table 9.

Table 9

**Significance of Group Membership**

Dependent Variable: Postest

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected Model</td>
<td>.479^a</td>
<td>2</td>
<td>.240</td>
<td>13.632</td>
<td>.000</td>
<td>.283</td>
</tr>
<tr>
<td>Intercept</td>
<td>2.628</td>
<td>1</td>
<td>2.628</td>
<td>149.506</td>
<td>.000</td>
<td>.684</td>
</tr>
<tr>
<td>Pretest</td>
<td>.157</td>
<td>1</td>
<td>.157</td>
<td>8.916</td>
<td>.004</td>
<td>.114</td>
</tr>
<tr>
<td>Group</td>
<td>.289</td>
<td>1</td>
<td>.289</td>
<td>16.455</td>
<td>.000</td>
<td>.193</td>
</tr>
<tr>
<td>Error</td>
<td>1.213</td>
<td>69</td>
<td>.018</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>38.201</td>
<td>72</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>1.692</td>
<td>71</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

^a. R Squared = .283 (Adjusted R Squared = .262)
The results from the ANCOVA show significance of group membership, $F(1, 71) = 16.45$, $p = 0.00$, demonstrating that students in Group 1 perform higher than students in Group 0 when controlling for pretest scores.

Table 10

<table>
<thead>
<tr>
<th>Model</th>
<th>R</th>
<th>R Square</th>
<th>Adjusted R</th>
<th>Std. Error of the Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.532a</td>
<td>.283</td>
<td>.262</td>
<td>.13259</td>
</tr>
</tbody>
</table>

a. Predictors: (Constant), Pretest, Group

A summary of the model shown in table 10 shows that 28% of the variance in student post-test scores is explained by pre-test score and group membership. The remaining 72% dependent on other variables unexplained.

Table 11

<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
<th>Correlations</th>
<th>Collinearity Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>Std. Error</td>
<td>Beta</td>
<td>T</td>
</tr>
<tr>
<td>1 (Constant)</td>
<td>.519</td>
<td>.049</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group</td>
<td>.128</td>
<td>.031</td>
<td>.415</td>
<td>4.056</td>
</tr>
<tr>
<td>Pretest</td>
<td>.489</td>
<td>.164</td>
<td>.305</td>
<td>2.986</td>
</tr>
</tbody>
</table>

a. Dependent Variable: Postest

The coefficients of the linear regression shown in Table 11 indicate that when moving from the control group to the experimental group, there is a 0.128 increase in post-test score when controlling for pre-test scores. Conversely, when increasing the pre-test score, there is a 0.489 increase in post-test score when controlling for group membership.
Additionally, tests to see if the data met the assumption of collinearity indicated that multicollinearity was not a concern also seen in Table 11 (Coefficients, Group, Tolerance = .99, VIF = 1.005; Pretest, Tolerance = .995, VIF = 1.005).

To help better understand the interaction between the variables, we would also like to test the effect of teacher in the study, therefore running a model with pretest, group, and teacher as well as the group by teacher interaction.

Table 12

<table>
<thead>
<tr>
<th>Group and Teacher Interaction</th>
<th>Dependent Variable: Postest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source</td>
<td>Type III Sum of Squares</td>
</tr>
<tr>
<td>Corrected Model</td>
<td>.743&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Intercept</td>
<td>2.862</td>
</tr>
<tr>
<td>Pretest</td>
<td>.069</td>
</tr>
<tr>
<td>Group</td>
<td>.230</td>
</tr>
<tr>
<td>Teacher</td>
<td>.144</td>
</tr>
<tr>
<td>Group * Teacher</td>
<td>.105</td>
</tr>
<tr>
<td>Error</td>
<td>.950</td>
</tr>
<tr>
<td>Total</td>
<td>38.201</td>
</tr>
<tr>
<td>Corrected Total</td>
<td>1.692</td>
</tr>
</tbody>
</table>

a. R Squared = .439 (Adjusted R Squared = .405)

From the table above, we can see that 43.9% of the variance is explained by these variables, leaving 66.1% unexplained (an improvement upon the previous model). We also notice a group by teacher interaction of $F(1, 71) = 7.384, p = 0.008$, indicating the effect of group on post test score differs by teacher.
Although scores increase in both cases, when we move from control to experimental group, there is a larger increase in score for teacher 2. Fidelity of implementation could play a role in this result. The graph below of the mean post test scores below shows this relationship.

Figure 1. Mean achievement variance of posttest scores by teacher
Qualitative Results

Focus group interviews were conducted with students and teachers to help better frame and understand the results from the quantitative portion of the study. In these semi-structured interviews, students were asked about their feelings toward mathematics, grading, and their experiences during the semester with the standards-based grading system used in their class. Recordings of the interviews were then transcribed, coded, and subsequently analyzed for thematic patterns and trends as seen in tables 14 and 15.

Table 14

<table>
<thead>
<tr>
<th>Investigated Themes</th>
<th>Construct</th>
<th>Frequency construct mentioned during interviews</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher feedback</td>
<td>Asking questions to help understand</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Teachers clarifying things in class</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Teachers providing specific feedback helpful</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Communication is important</td>
<td>9</td>
</tr>
<tr>
<td>Standards-based grading</td>
<td>Presenting mastery information in a clear way</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Providing specific and focused information</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Helping guide remediation and assist teachers</td>
<td>10</td>
</tr>
<tr>
<td>Student Achievement</td>
<td>Aided by standards-based grading and feedback</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Confusion with traditional grades</td>
<td>7</td>
</tr>
</tbody>
</table>

As evidenced by the constructs coded above in Table 14, students responded in an overwhelmingly positive fashion toward the adoption of standards-based grades in their mathematics class. Overarching themes consistent with the new grading system were clarity,
specificity of feedback, and helpfulness toward students and teachers when focused upon improvement.

Table 1

<table>
<thead>
<tr>
<th>Investigated Themes</th>
<th>Construct</th>
<th>Frequency construct mentioned during interviews</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student engagement</td>
<td>Making math relevant to students lives</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Connecting all of the pieces together</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Establishing and maintaining student relationships</td>
<td>9</td>
</tr>
<tr>
<td>Feedback for Students</td>
<td>Importance of regular and specific feedback</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Informal assessment through questioning</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Using data from tests to help inform instruction</td>
<td>8</td>
</tr>
<tr>
<td>Standards-based grading</td>
<td>Aided by standards-based grading and feedback</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Helps with remediation and improvement</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Takes more time to maintain</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>More clear than traditional grading</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Tests were longer</td>
<td>3</td>
</tr>
</tbody>
</table>

Teachers responded in a mostly positive fashion when asked about their experiences with grading and its affect on student achievement during the semester, as evidenced by the frequency of responses in Table 15. Overarching themes consistent throughout the interviews were the importance of building relationships with students, providing feedback and encouragement, and using data from assessments to help guide instruction and remediation. Some negative aspects of the new grading systems were noted as teachers recognizing the additional commitment in time.
needed to employ the detailed rubrics as needed, and the additional length of the tests needed to
gauge all skill masteries within a unit.

**Summary of Quantitative and Qualitative Data**

This research study employed a quantitative data component as well as qualitative data
compontent in an effort to better answer the questions framed by the researcher:

1. Does the way in which students are graded exhibit a measurable influence upon
   overall student achievement in a secondary mathematics class?
2. To what extent do students perceive grading systems as having influence on their
   achievement in secondary mathematics?

The researcher used an ANCOVA test to analyze the significance of group membership
(experimental standards-based grading or control traditional grading) on student achievement as
measured by a pre-test and post-test for students enrolled in North Carolina Math 1. The findings
shown in table 9 show a statistically significant relationship exists for students in the
experimental group and increased achievement ($\eta^2 = 0.193$). Homogeneity grouping and slope
tests were conducted to ensure the pre-conditions for an ANOCVA were met (Table 3, Table 8).

Additionally, qualitative focus group interviews were held to help better frame and
understand the quantitative portion of the study. Findings from a thematic analysis of both
interview groups results in a generally positive feeling toward standards-based grading with
emergent themes focusing upon its clarity in feedback, helpfulness to teachers and students, and
specific and targeted information regarding student achievement. Some negative aspects noted
included the additional time required to implement a system as well as the perception that
standards-based assessments were longer and took more time to conduct.

Generally, the quantitative results shown above indicate a positive affect of the
implementation of standards-based grading with students demonstrating higher levels of achievement than in the traditionally graded classrooms. These results in students test scores are clarified and supported by student responses given in the focus group sessions. Student responses such as “standards clearly let you see what you need to improve upon” and that the standards-based feedback given by teachers was “detailed” and what “made a teacher great.” Teachers also responded in a similar fashion in that it “helped guide remediation and review” and was useful for “self-reflecting” on what students knew and didn’t know.
Chapter 5: Discussion

Introduction

The purpose of this case study was to determine whether or not student grading paradigms and methods have a significant impact upon the level of student achievement in secondary mathematics, as well as students’ perception of their own achievement in those courses. While many similar studies exist that focus upon pedagogy and instructional techniques, little to no research has been done into the effect that a detailed rubric-based mechanism of student feedback may have upon student achievement in secondary mathematics. While rubrics and standards-based grading have been proven useful at the elementary level (Gusky, Swan, & Jung, 2011), their implementation at a secondary level is often inhibited by complexity of curriculum standards and lack of teacher familiarity and subsequent buy-in.

A review of the literature suggests that math students in the United States are lagging behind their world counterparts significantly (Silver, 2014) which in turn places them at a competitive disadvantage when attempting to be competitive in a global job marketplace (Hossain & Robinson, 2012). These deficiencies have led educational reformers to strive for both innovative and carbon copy solutions to help bridge the gap to help better prepare students for their future. Teacher behavior and practices are seen to have the largest impact upon student achievement (Goddard, Hoy & Hoy, 2000) and thus naturally have become the primary focus of educational reforms and accountability measures.

Rubrics are used by teachers and instructors throughout various levels of education to help provide clear outlines of learning objectives as well as detailed feedback that helps inform the learner as to their progress toward mastery (Jonsson and Svingby, 2007). This practice aligns with research by Marzano (2007) that demonstrates 29 and 23 percent gains when reinforcing
effort on an individual basis and setting specific objectives then subsequently providing detailed feedback, respectively. It was the researcher’s hope that this study would help clarify and better understand the effects of standards-based grading, driven by the use of detailed rubrics, upon secondary math achievement. Additionally, this study will also contribute to the literature and advance the discussion of grading reform in the secondary classroom. The discussion of the findings will be followed by implications for the educational field, observed limitations of the findings, and recommendations for additional research.

**Findings**

The research questions used in this study to determine the effect of standards-based grading in the secondary mathematics classroom are as follows:

1. Does the way in which students are graded exhibit a measurable influence upon overall student achievement in a secondary mathematics class?

2. To what extent do students perceive grading systems as having influence on their achievement in secondary mathematics?

The case study used to help answer these questions employed both quantitative and qualitative data collection to gain a better understanding of the results. The experimental group of students ($n = 33$) were exposed to standards-based grading throughout the duration of the semester, while the control group of students ($n = 39$) were evaluated using traditional grading systems.

The quantitative findings present several pieces of information in support of standards-based grading. First, the control classes as a whole experienced less growth on average ($M = 39.56\%, \ SD = 15.15\%$) than the experimental classes ($M = 50.51\%, \ SD = 14.96\%$). While this could be possibly attributed to student characteristics and math abilities, a one-way ANOVA
demonstrated there is no statistically significant effect from membership in either group, $F(1,72) = 0.37, p = 0.55$ while a regression test also showed that the data sets were parallel when the interaction term was found to not be insignificant ($F(1,71) = 0.106, p = 0.745$).

When controlling for pre-test scores, the results from the ANCOVA show significance of group membership, $F(1, 71) = 16.45, p = 0.00$, demonstrating that students in the experimental classes perform higher than students in the control classes. This finding seems to help answer the researcher’s first research question, indicating that students being graded using a standards-based mechanism throughout the semester experienced higher levels of achievement and growth throughout the semester. Similarly, the research points to a grading methods as a sizeable influence upon student achievement, with 28% of the variance in post-test score being attributed to group membership.

To better frame the results of the quantitative portion of the study, qualitative focus group interviews were conducted in an attempt to help answer the researcher’s second research question regarding student perception of their achievement when using standards-based grading. It was the researcher’s hope that integrating student and teacher feedback with the quantative data obtained previously would provide a more complete picture and accurate perception of the impact and generalizability of the findings (Gelo, Braakmann & Benetka, 2008). Student perception of the effectiveness of feedback received from teachers largely impacts their self-efficacy (Bandura, 1977). This personalization hopefully leads to more meaningful teacher relationships in which mutual learning and respect create an environment conducive to success, in alignment with Bandura’s (1977) research on social cognitive theory.

Students and teachers interviewed for the study responded positively toward the use of standards-based grading in their mathematics class. Themes emerging from the discussion
support the findings in the qualitative portion of the study in that students found themselves to be “better prepared” for summative assessments and that teacher feedback was “more detailed” and helped to “clarify” what students knew and did not yet know. Students also appreciated the organization of the standards-based rubrics in that “it showed [the student] what [they] knew and didn’t know” whereas a traditional percentage and letter grade experienced previously led to students being “not really sure” where they excelled and fell short. Teachers similarly found to be beneficial the targeted feedback in that it helped them prepare for remediation and plan for additional practice to help students grow toward mastery.

Interestingly, there was found to be a difference in the effect of the treatment between the two different teachers participating in the study. Teacher 1’s students showed less of an achievement gain when moving from the control group to the experimental group than Teacher 2. Reasonable explanation for this would stem from a difference in the fidelity of implementation. Additionally, controlling for all other variables, the control group for Teacher 2 had the lowest level of achievement overall. It would be interesting to re-interview teachers that participated in this study for a future exploration of factors affecting the variance in outcomes.

**Implications**

Students participating in this study, as well as students throughout the country are in need of improved mathematics skills and achievement to better prepare themselves for post-secondary study. Student access and degree completion at the post-secondary level is one of largest factors in marketability toward prospective employers (Sparks & Waits, 2011). Better preparing those students during their K-12 studies can help those students be more successful in their post-secondary studies. A lack of preparation can lead to students not seeking a college education, or
Students who experience academic success are more likely to secure rewarding employment and make a significant contribution to society. It naturally follows that educational reforms focused upon improving student academic success can largely impact society as a whole. Transformative educational policy aimed at increasing accountability and teacher effectiveness have been a large part of the landscape over the past several decades. The achievement gap between American students and their counterparts throughout the world is apparent, recognized, and accepted as fact. Policy makers strive to find solutions to address this problem from various angles, including increased resources, increased testing and accountability, and standardization of curriculum. While these initiative do not always align with educational research findings, they are targeted at improvement of a teacher’s practices, the largest quantifiable influence upon student achievement (Darling-Hammond, 2000).

Standards-based grading mirrors the curriculum standards that students are measured against. It provides a method of feedback that is detailed, specific, and personalized to the student in alignment with research that shows these characteristics of teacher behavior to be beneficial to overall student learning. Additionally, it provides clarity in understanding what students have mastered. If the overall goal of education is for students to achieve mastery, a more detailed measuring instrument to evaluate that achievement would be beneficial to all of those involved. Employing such a practice allows stakeholders – parents, students, teachers, and administrators – to be more involved and have a better understanding of student success. This creates the opportunity for those involved to then use that understanding to help develop additional targeted interventions to help support students as they work toward overall mastery.
The results of this study indicate that students benefit from standards-based grading in the secondary mathematics classroom. When teachers employed this method of grading in their classroom during this study, students were found to achieve higher levels of mastery, as well as feel more informed about their performance and abilities. Student’s also appreciated the level of personalization and the fostering of the relationship between the teacher and student that followed. The quantitative portion of this research study indicates a significant relationship between standards-based grading and higher levels of student achievement. When taken in conjunction with the qualitative interviews of students and teachers, these findings support the idea that standards-based grading is beneficial to students and teachers in the secondary mathematics classroom.

This is not to say that a transition to a standards-based grading system in secondary mathematics would be an easy one. A significant amount of time was invested in developing standards-based targets and feedback instruments to employ throughout the semester. Additionally, a great deal of information and discussion were required to help students and parents understand the new grading procedures and policies implemented in class. However, it is the opinion of the researcher that this should not be seen as a deterrent. While a great deal of time was invested in the creation and implementation of standards-based grading in the Math 1 classroom this time would be greatly reduced in future semesters by using what has already been created and revising it as necessary. Any changes in curriculum goals could be reflected through a smaller and more manageable revision process. Widespread adoption of standards-based grading and its procedures would also better inform and prepare students and parents, simplifying the acclimation to its use.
Limitations

This study had several limitations that the researcher wishes to acknowledge. First, the sample sizes in the study were smaller than ultimately desired. A small sample size can lead to possible ambiguity in the interpretation of quantitative results. A larger sample size at a larger school would help to solidify the results experienced in this study.

Additionally, the sampling method used in this study, convenience sampling, can sometimes lead to inherent bias that can make generalization of the results somewhat difficult. This method of sampling can also lead to under-representation of over-representation of particular groups within the sample. Truly random sampling was not feasible in this study due to limitations in student and teacher scheduling.

The location of this study and the students it involved are also confined to one location and school setting. This setting may have unique demographics or properties that can make generalization of the results to a broader setting problematic.

Lastly, this body of research would have benefited greatly from additional time to study the longitudinal effects of standards-based grading on the students sampled from the population. Doing so would have given a more definitive picture of its effects with multiple teachers and curriculums throughout secondary mathematics.

**Recommendation for Future Research**

It is the researcher’s hope that this study provides ample foundation for future research aimed at answering similar research questions, but on a larger scale. Student achievement continues to and will always be the elusive moving target at which educational researchers aim to successfully hit with reforms and initiatives. Standards-based grading can be seen as yet another useful aramement in the battle to better prepare students working toward mastery.
Future research could focus upon different forms and implementation of standards-based grading in an effort to better discover the qualities that make it most effective. It would also be beneficial to attempt to transpose the findings in this study from secondary mathematics to other content areas in secondary education. The literature suggests that standards-based grading is more prevalent in elementary settings. Longitudinal studies that examine student achievement while being graded with standards-based grading throughout the duration of the K-12 years would be helpful in understanding its long-term impacts. Additionally, conducting similar studies in various settings and with various demographics would also be helpful in understanding its effects.

Given the sizeable difference in treatment effect between the two participating teachers, it would be beneficial to conduct additional research surrounding the circumstances of implementation. Differences in the approach of the teacher, communication with the students, and other factors not examined in this study could help further clarify best practices for implementation.

**Conclusions**

This study aimed to better understand the way in which student grading affects student achievement and perception of achievement. By examining student achievement in secondary mathematics classes throughout the duration of a semester, the study aimed to quantify the impact that standards-based grading had upon student achievement and growth. Students who experience a lack of academic success often encounter obstacles that make it difficult to prosper as a member of society. Additionally, a lack of academic success can translate into behavioral and emotional difficulties, results in a lack of self-efficacy and a feeling of general lack of worth (Zajacova, Lynch & Espenshade, 2005). The results of this study show that providing students
with personalized, detailed feedback using standards-based grading can be seen to improve student achievement.

It should be noted that prior to implementing standards-based grading at the secondary level, a great deal of prior planning and involvement of teachers should be employed to create an environment conducive to its success. Teachers should be directly involved with developing the learning targets that evaluate mastery in alignment with curriculum standards. Creative freedom with the tracking and feedback mechanisms of standards-based grading will help teachers adapt it to their already existing practices. Standards-based grading is not a miracle cure to the problems that ail American students. It is, however, based upon this research study, a purposeful and useful ingredient in the overall treatment plan.
References


Marzano, R. J. *Designing & Teaching Learning Goals & Objectives*.


Education, Inc.


Curriculum, College Graduation, and Earnings (pp. 1–178).


Schmoker, M., & Marzano, R. J. (1999). Realizing the promise of standards-based


achievement in high school students. Psychological reports, 98(2), 318-322.

Shippy, N., Washer, B. A., & Perrin, B. (2013). Teaching with the end in mind: The Role of

Instructional Guidance in Mathematics' Mathematics and Authentic Assessment.

Silver, E. A. (2014). Using PISA to Stimulate STEM Teacher Professional Learning in
the United States: The Case of Mathematics. Issues in Teacher Education,
23(1), 11-30.


Appendix A
Standards-Based Assessment Exemplar

Standard 1.01 – Use order of operations to evaluate expressions
½  1 Evaluate expressions using the correct order
½  1 Evaluate exponents correctly and in the right order
½  1 Substitute and evaluate expressions using the correct order
½  1 Evaluate expressions with negative numbers correctly

Proficiency Score  ____ out of 4

1) $8 - 2 / 1 + 4 * 3$  2) $2(3 + 2) - 4 / 2$  3) $2 - 4(3 + 1)^2$
4) $(4 - 2)^2 + 5$  5) $5(2)^3 + 12 / 3 + 5$  6) $\frac{8-(2+3)}{4+3}$

Given that $x = 3$, $y = -2$, and $z = 4$, evaluate each of these expressions.

7) $2x + 8y$  8) $3x^2 - 3x + 2$  9) $yyy^2$  10) $\frac{xy}{x+2(yz)}$

Standard 1.02 – Use variables and expressions to represent common language
½  1 Translate from English to mathematics
½  1 Solve one-step equations correctly
½  1 Combine like terms correctly
½  1 Express and rate of change and starting point in equation form

Proficiency Score  ____ out of 4

Write an expression for the verbal statement. Then solve for the value of the variable.

11) Seven more than half a number $x$ is 9.  12) 9 less than a number $y$ squared is 7.

13) A quarter of a number $a$ increased by 2 is 3.  14) The product of $x$ and 3 is 21.

Combine like terms and give a variable expression for each figure’s perimeter.

15) $2x - 4 + x + 7 = 4x + 3$  16) $2x + 5 = \frac{3x + 1}{2}$
17) My cell phone plan charges me a $50 access fee per month for unlimited calling and texting, and $15 per GB of data I use.

   a) Calculate my bill for 2 months after I make 100 calls, send 30 text messages, and use 8 GB during the 2 month period.

18) A toll road in Chicago charges $0.35 for every 10 miles traveled on the road.

   b) East Chicago is 40 miles from West Chicago. Calculate my toll.

19) An ice machine costs $300 to rent for the summer, and the profit made off each bag of ice is $2.00.

   a) How many bags of ice do I need to sell to break even (not lose money)?

<table>
<thead>
<tr>
<th>Standard 1.03 – Solve linear equations with one or more steps.</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \frac{1}{2} ) 1 Solve one-step equations.</td>
</tr>
<tr>
<td>( \frac{1}{2} ) 1 Solve multi-step equations.</td>
</tr>
<tr>
<td>( \frac{1}{2} ) 1 Use distributive and like terms correctly.</td>
</tr>
<tr>
<td>( \frac{1}{2} ) 1 Work effectively with negative numbers.</td>
</tr>
</tbody>
</table>

Proficiency Score _____ out of 4

Solve each equation for the value of \( x \).

20) \( x - 12 = 30 \)  \hspace{1cm} 21) \( -2x = -11 \)  \hspace{1cm} 22) \( \frac{1}{2}x = 24 \)
23) \( 5x - 4 = 3x - 6 \)  \hspace{1cm} 24) \( 4x - 10 = -2x + 8 \)  \hspace{1cm} 25) \( 5x + 3 = 2(x + 6) \)
26) \( 3x + 6 - x = x - 14 \)  \hspace{1cm} 27) \( 21 + 3(x - 4) = 4(x + 5) \)  \hspace{1cm} 28) \( 12 - \frac{1}{4}x = 5 \)

<table>
<thead>
<tr>
<th>Standard 1.04 – Solve literal equations for indicated variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \frac{1}{2} ) 1 Solve formulas with one step</td>
</tr>
<tr>
<td>( \frac{1}{2} ) 1 Solve multi-step formulas</td>
</tr>
<tr>
<td>( \frac{1}{2} ) 1 Translate from English to math and then rearrange</td>
</tr>
<tr>
<td>( \frac{1}{2} ) 1 Solve for the correct variable</td>
</tr>
</tbody>
</table>

Proficiency Score _____ out of 4

Rearrange each of the following formulas for the indicated variable.

29) \( E = IR; \) solve for \( I \).  \hspace{1cm} 30) \( C = 2\pi r; \) solve for \( r \).  \hspace{1cm} 31) \( P = 2a + b; \) solve for \( b \).
32) \( T = \frac{X + Y + Z}{3}; \) solve for \( X \).  \hspace{1cm} 33) \( y = mx + b; \) solve for \( m \).
34) The number of girls (G) in a class is 5 more than twice the number of boys (B). Write an equation that is solved for $B$.

35) The number of cars (C) in a parking lot is 3 less than three times the number of trucks (T). Write an equation that is solved for $T$.

<table>
<thead>
<tr>
<th>Standard 1.05 – Solve and graph linear inequalities</th>
</tr>
</thead>
<tbody>
<tr>
<td>½</td>
</tr>
<tr>
<td>½</td>
</tr>
<tr>
<td>½</td>
</tr>
<tr>
<td>½</td>
</tr>
</tbody>
</table>

**Proficiency Score ___ out of 4**

Solve each inequality for the value of $x$ and then graph the solution on a number line.

36) $2x \geq -12$  

37) $x - 4 \leq 18$  

38) $12 - x > 7$

39) $4x + 11 \geq 1$

40) $16 + 2x > 4x - 1$

41) $3(x - 3) < 12 + x$
Appendix B

Standards-Based Feedback Exemplar

Standard 1: Operations with Numbers, Equations, and Inequalities

- 1.01 – Use order of operations to evaluate expressions
  - a – integers only
  - b – substitution with variables
    ______ out of 4

- 1.02 – Use variables and expressions to represent common language
  - a – translate words to math
  - b – combine like terms
  - c – solve simple equations from translations
    ______ out of 4

- 1.03 – Solve linear equations with one or two steps
  - a – one step equations
  - b – two step equations
  - c – variables on both sides
  - d – distributive, like terms, fraction bars
    ______ out of 4

- 1.04 – Solve literal equations for indicated variables
  - a – rearrange with one-step
  - b – rearrange with two-steps
  - c – translate and then rearrange
    ______ out of 4

- 1.05 – Solve and graph linear inequalities
  - a – one step inequalities
  - b – two step inequalities
  - c – graphing solutions on a number line
    ______ out of 4