THE IMPACT OF STANDARDS-BASED GRADING ON THE ACADEMIC SELF-EFFICACY OF HIGH SCHOOL STUDENTS: A MIXED METHODS STUDY

A Dissertation

Presented in Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

With a

Major in Educational Leadership in the

Department of Graduate Education

Northwest Nazarene University

by

Derek Bub

April 2021

Major Professor: Dr. Bethani Studebaker, Ed.D

DocuSign Envelope ID: EAD8A8A6-5F0D-4EC5-8248-9CDBABC80663

AUTHORIZATION TO SUBMIT

DISSERTATION

This dissertation of Derek Bub, submitted for the degree of Doctor of Philosophy with a major in Educational Leadership and titled "The Impact of Standards-Based Grading on the Academic Self-Efficacy of High School Students: A Mixed Methods Study," has been reviewed in final form. Permission, as indicated by the signatures and dates given below, is now granted to submit final copies.

Major Professor DocuSigned by: Bethani Studebaker DFE07B692FFB4DB... Dr. Bethani Studebaker Date 5/3/2021 | 09:48:28 MDT

Committee Members DocuSigned by: Dr. Guy Kovacs F4678E8A95F843F... Dr. Guy Kovacs Date 5/3/2021 | 13:58:49 MDT

DocuSigned by: C9664850D8C0460... Dr. Eian Harm Date 5/4/2021 | 07:42:25 MDT

Doctoral Program Director DocuSigned by: Bethani Studebaker DFE07B692FFB4DB... Dr. Bethani Studebaker Date 5/3/2021 | 09:48:28 MDT

Discipline's College Dean DocuSigned by: Lori Sanchez 1F6287564ACC4DC... Dr. LoriAnn Sanchez Date 5/4/2021 | 09:57:38 MDT

© Copyright by Derek Bub 2021

All Rights Reserved

ACKNOWLEDGMENTS

First and foremost, I would like to thank God for providing me with the talents and abilities He has gifted me with to pursue and accomplish this goal. My prayer is that I continue to find ways to use these gifts to leave a lasting impression on this world and His children. I would also like to thank Him for the clear direction throughout this process and for the extra push every time my tank was empty. I strongly believe that His gift to me are the talents and abilities He has given me, and my gift to Him is what I do with them.

Second, I would like to thank Dr. Studebaker for her unending support and guidance throughout this process. I can honestly say that this would not be possible without her unwavering commitment to my success. She has been a pillar, a sounding board, and an experienced leader on this entire journey. Her willingness to go above and beyond, time and time again, and take afterhours phone calls to provide words of encouragement or clear direction were an integral part of this process. She continually challenged me and pushed me to be a better educational leader each and every day, for which I am sincerely and eternally grateful.

To my committee members, Dr. Harm and Dr. Kovacs, thank you for your commitment and for investing your time into this endeavor. I sincerely appreciate your willingness to serve on this committee and be a critical part of the process. Your dedication provided me with the opportunity to set and accomplish this goal. Without you, the journey would simply not have been the same, and I thank you.

To my colleagues, both current and former, thank you for believing in me. I have been lucky enough to come across so many people in my professional career that have believed in me and challenged me to continue my education. These words of encouragement have given me the confidence and faith that all things are possible.

DEDICATION

I recognize that I am not in my present position because of a simple act of luck or special circumstance. I am fortunate to be surrounded by people that consistently inspire me and provide support as I work to make a positive impact in this world. Without each of these people, this accomplishment would not be possible.

My amazing wife Becky, who lovingly refers to life with me as a rollercoaster: you are not just a willing participant on this ride, but also a foundational support every time I look to pursue another incredible adventure. You continually help me take the impossible and make it possible, and I thank you for being my teammate, pillar, and confidante in life. My daughters Delaney and Savannah: your patience and understanding as we have walked this journey has been tremendous. The memories of completing annotations during a volleyball game will not be forgotten. Thank you for your unconditional love and encouragement throughout this process.

My Dad, who is no longer with us on earth, but I know he is watching over me right now: I hope I have made you proud. It is the work ethic that you instilled in me throughout my life that has allowed me to accomplish this goal. I did not always recognize it growing up, but it is long overdue for me to say, thank you, and I love you Dad.

My friends and mentors who have provided me with the network that has carried me through the actualization of this goal. Your compassion and understanding throughout this journey made this experience possible. Your motivation and guidance kept me focused and determined to reach the finish line. I can honestly say I would not have made it without your support. In reflection, I recognize that I am a product of those that I have been lucky enough to encounter in my life, I am incredibly thankful for the people around me, both past and present, that have invested time and energy into allowing this journey to come to a completion.

ABSTRACT

In recent years, several notable educational scholars have critiqued the traditional way that the grading of students is completed in the United States. This criticism has led many educational experts and institutions to consider reforming grading practices. One of the solutions that several districts have attempted to implement is referred to as standards-based grading. This form of grading focuses on reporting student progress specific to their achievement on a predetermined standard. The purpose of this study was to analyze the impact of a standards-based grading philosophy on the academic self-efficacy in English classes of juniors in high school. This study looked at students from two different high schools, located in the Pacific Northwest region, where a standards-based grading approach had recently been implemented. In a pre then post retrospective design, students were asked to reflect on their self-efficacy prior to the standards-based grading model and their current self-efficacy in the new standards-based grading model. The results of the surveys were disaggregated, and semi-structured interviews were conducted with volunteer participants to provide deeper understanding of the survey results. The Wilcoxon Signed-Rank test was used to analyze the correlation between self-efficacy scores before and after the implementation of standards-based grading. This analysis showed a positive correlation after the implementation of standards-based grading. Small group, semi-structured interviews were then conducted, recorded, and transcribed. The transcriptions of these interviews were coded and themed to examine the influence of this new grading system on the academic self-efficacy of students. Qualitative and quantitative results indicate that standards-based grading has a positive impact on the academic self-efficacy in English classes of high school juniors.

TABLE OF CONTENTS

ACKNOWLEDGMENTS ii
DEDICATION iii
ABSTRACT iv
Chapter I Introduction 1
Statement of Problem 7
Background 10
Research Questions 13
Definition of Key Terms 14
Significance of the Study 16
Theoretical Framework 16
Overview of Research Methods 17
Chapter II Review of Literature 19
Introduction 19
Theoretical Framework 21
History of Traditional Grading Practices 21
Deficiencies of Traditional Grading Practices 30
Rationale for Standards-Based Grading 33
Implementation of Standards-Based Grading 37
Standards-Based Grading for Parents 39
Standards-Based Grading for Students 43
Standards-Based Grading and Student Achievement 45
Student Perceptions of Fairness and Academic Achievement 46
Conclusion 46
Chapter III Design and Methodology 49
Introduction 49
Research Questions 50
Research Design 50
Data Collection 57
Quantitative Data 59
Qualitative Data 60
Analytical Methods 61
Limitations 61
Chapter IV Results 63

Introduction 63
Participant Profile 65
Vignettes of Schools 66
Participant Profile 67
Interview Participants 70
Survey Validity and Reliability 71
Semi-Structured Focus Group Pilot Validity and Reliability 73
Data Collection Instruments 74
Quantitative Results 76
Quantitative Results for Question #1 77
Qualitative Results for Question #2 82
Qualitative Research Findings 86
Discussion of Emerging Themes and Findings 89
Conclusion 94
Chapter V Discussion 95
Introduction 95
Summary of Results 98
Research Question #1: Summary of Results and Discussion 100
Research Question #2: Summary of Results and Discussion 105
Conclusions 107
Recommendations for Further Research 110
Implications for Further Practice 111
References 116
Appendix A General Self Efficacy Scale 128
Appendix B Adapted Survey 130
Appendix C Informed Consent 133
Appendix D Clinical Research Ethics Certification 136
Appendix E Institutional Review Board Approval 137
Appendix F Site Approval Letters 138
Appendix G Board Approval Letter 141
Appendix H Qualitative Interview Questions 142
Appendix I E-mail to Teachers Requesting to Use Class for Study 143
Appendix J NNU Interview Protocol 144
Appendix K Permission for Tables 145

Appendix L Permission for Tables 146
Appendix M Permission for Tables 147

LIST OF TABLES

Table 1 *Comparison of Traditional Grading and Standards-Based Grading* 36
Table 2 *Sample Four-Year Standards-Based Grading Implementation Plan* 38
Table 3 *Participating Schools Demographic Breakdown* 54
Table 4 *Participating School* 56
Table 5 *Participants for Interviews* 57
Table 6 *Survey Participants and Interview Volunteers* 59
Table 7 *Participant Demographics by Location (n=63)* 69
Table 8 *Focus Group Participant Demographics* 71
Table 9 *Cronbach's Alpha Reliability* 73
Table 10 *Demographics of Pilot Interview Participants* 74
Table 11 *Impact of Standards-Based Grading on Self-Efficacy of Students* 78
Table 12 *Survey Results for Research Question 1 by Gender* 80
Table 13 *Survey Results by Question* 84
Table 14 *Key Words for Qualitative Themes* 88
Table 15 *Frequencies of Themes* 89
Table 16 *Rank by Time in SBG* 101

LIST OF FIGURES

Figure 1 *Traditional Grading System vs. Standards-based Grading System* 4
Figure 2 *Categories of the Literature Review* 21
Figure 3 *Social Cognitive Theory-Process of Goal Realization* 22
Figure 4 *Determining Efficacy Judgements* 24
Figure 5 *Example: Standards-Based Grading Report Card* 42
Figure 6 *Brunswick Academy Report Card* 43
Figure 7 *Interconnected Relationship between a Grading System, Self-Efficacy, Persistence, and Task Selection* 46
Figure 8 *Conceptual Process of Explanatory Sequential Mixed-Methods Design* 53
Figure 9 *Percentage of Positive Ranks* 82
Figure 10 *Word Frequency* 87
Figure 11 *Influences of Self-Efficacy* 102

Chapter I

Introduction

Douglas Reeves (2008) stated, "If you wanted to make just one change that would immediately reduce student failure rates, then the most effective place to start would be challenging prevailing grading practices" (p.85). Over the past four decades, educational reform has focused on assigning standards to be taught in each grade level and subject area with very little movement in grading practices (Guskey, 2011; Townsley & Buckmiller, 2016). The adoption and implementation of the Common Core standards in many states have made learning targets clearer and more rigorous for all students (Davis, 2019). However, reforms in grading practices that focus on the most effective ways to assess student achievement on specific standards, and correlating the achievement to accountability measures, have been largely untouched (Guskey, 2011; Townsley & Buckmiller, 2016). The purpose of grading is to report on student progress toward specific learning goals related to a class and to communicate this information to all stakeholders (Anderson, 2018;; Guskey & Munoz, 2015; O'Connor, 2007; Swan, Guskey, & Jung 2014). However, the validity of the traditional grading system has been challenged and shown to be an unreliable source of information to parents, students, teachers and educational leaders (Marzano, 2000; O'Connor, 2007; Schimmer, 2016).

In traditional grading systems, there is a wide variety of factors that influence a grade on a particular assignment or the final outcome in a class. The wide variety of variables that impact grading in traditional systems lead stakeholders to feel ill-informed about what a specific grade communicates towards a student's overall academic progress (Brookhart, 2009; Guskey & Munoz, 2015; Marzano, 2000; O'Connor, 2007; Schimmer, 2016). In this grading system, a

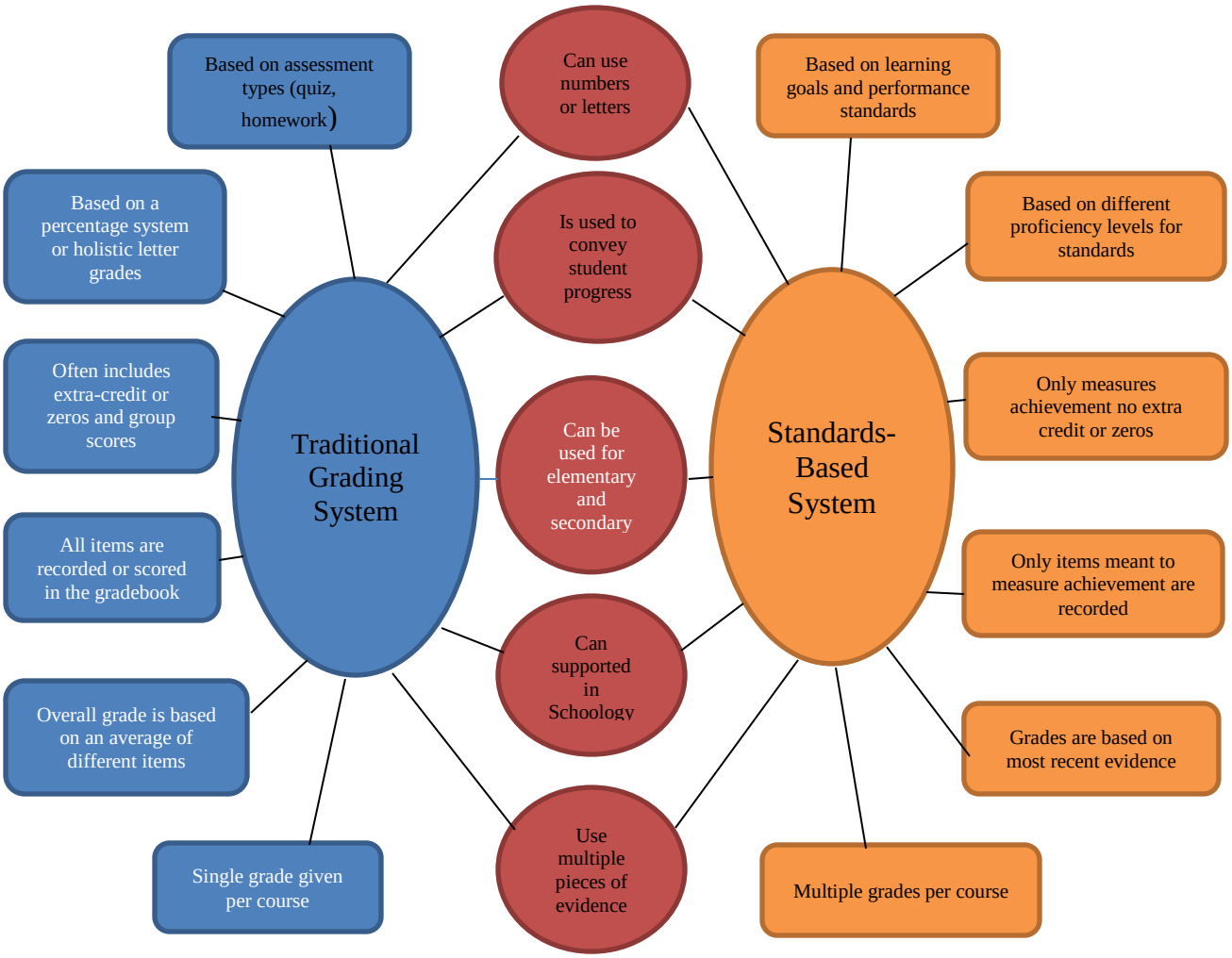
teacher determines a grade (A, B, C, D, or F) based on a student's total achievement in that class; this includes an evaluation on various activities, assignments, and behaviors (Heflebower & Marzano, 2011; Marzano, 2000; O'Connor, 2007; Schimmer, 2016). All of these factors are weighted and averaged to help determine a final percentage and overall grade, as a mechanism to sort students from highest to lowest (Heflebower & Marzano, 2011; Marzano, 2000; O'Connor, 2007; Schimmer, 2016). Non-academic and other invalid measures are also included in a student's grade, which leads to miscommunication regarding a student's academic performance for a specific subject area (Heflebower & Marzano, 2011; Marzano, 2000; O'Connor, 2007; Schimmer, 2016). Behavioral influences on a grade, such as extra-credit, points for timeliness, and zeros for incomplete work, can create inaccurate and confusing reporting for both student and parents (Brookhart, 2009; Heflebower & Marzano, 2011; O'Connor, 2007; Schimmer, 2016). Research indicates that the traditional grading system yields very little useful information about a student's academic progress (Anderson, 2018; Heflebower & Marzano, 2011; Marzano, 2000; O'Connor, 2007; Schimmer, 2016). A grade of a "B" in a class may indicate that the student has outstanding behavior but very little content knowledge. In contrast, a grade of a "D" in a class could represent that the student has poor behavior with outstanding content knowledge (Heflebower & Marzano, 2011; Marzano, 2000; O'Connor, 2007; Schimmer, 2016).

The lack of clarity provided in traditional grading systems has forced teachers and districts to look at reforming current grading practices into standards-based grading in an effort to better communicate student achievement specific to each standard (Brookhart, 2011; Guskey, 2011; Heflebower & Marzano, 2011; Marzano, 2000; O'Connor, 2007; Schimmer, 2016). Standards-based grading, as defined by Robert Marzano (2010), is a system of grading and assessing a student's competency on specific standards within a subject area. In an effort to

minimize some of the pitfalls found in traditional grading systems, educators, parents, and students, have been rethinking long-held beliefs about grading (Brookhart, 2011; Heflebower & Marzano, 2011; Heflebower et al., 2014). The objective of standards-based grading, when properly implemented, is to take away the influences of behavior and other invalid measures to allow accurate reporting and communication on a student's achievement with regard to a particular standard (Brookhart, 2009; Heflebower & Marzano, 2011; Heflebower et al., 2014; O'Connor, 2007; Schimmer, 2016). This method of reporting improves communication about a student's academic achievement, explicitly helping stakeholders differentiate between learning in the process and the quality of student work; that is, understanding exactly what a student is working on, what they know, and their level of proficiency on the specific standard (Brookhart, 2009; Guskey, 2001; Heflebower & Marzano, 2011; Schimmer, 2016). Figure 1 (shown below) identifies the similarities and differences that are presented when comparing traditional grading practices and standards-based grading practices.

Figure 1.

Traditional Grading System vs. Standards-Based Grading System



Note. Adapted from Davis, L. (2020). Standards based grading: What to know in 2020.

Retrieved from: https://www.schoology.com/blog/standards-based-grading

For the purposes of this research, standards-based grading and standards-referenced grading will be used interchangeably. Although the intent of this study is to look at standards-referenced grading, standards-based grading is more commonly used in educational settings and publications (Heflebower et al., 2014; Marzano, 2010). These two grading concepts carry similar frameworks, with the critical difference between the two systems being that students in a standards-based system can immediately move to a more advanced standard once they have mastered the standard they are currently working on (Heflebower et al., 2014). "The vast majority of schools and districts that claim to have standards-based systems in fact have standards-referenced systems" (Marzano, 2010, p.18-19).

As many school districts across the United States look to reform grading practices in an effort to give more precise feedback regarding student learning, several are choosing to implement practices based on the standards-based grading philosophy (Welsh et al., 2013; Heflebower & Marzano, 2011; Munoz & Guskey, 2015). Educational reform does not always come easy, and transitioning a class, school, or district from a traditional grading system to a standards-based system can create some growing pains for leaders, parents, students, and teachers alike (Guskey, 2015; Heflebower et al., 2014; O'Connor, 2007; Swan, Guskey, & Jung, 2014). As schools attempt to make transitions into a standards-based grading system, they are often faced with resistance from stakeholders who are accustomed to traditional grade reporting (Franklin, Buckmiller, & Kruse, 2016; Peters & Buckmiller, 2015; Schimmer, 2016; Swan et al., 2014).

Making the transition from traditional grading practices to standards-based grading practices must be done in small, incremental steps (Schimmer, 2016). It is vital to create a standards-based mindset in all stakeholders before moving into a standards-based grading system

(Schimmer, 2016). Student input is also a critical factor in the effective implementation of a standards-based grading system. Schools looking to improve student grade reporting through the use of standards-based grading must include students in those conversations since they are the individuals most directly impacted by these decisions (Guskey et al., 2014; Marzano, 2000; Spencer, 2012). Getting other stakeholders to believe in this philosophy will ensure progress during challenging times (Schimmer, 2016). Parents, teachers, board members, and administrators must be shown the deficiencies of traditional grading practices and be offered the solutions that standards-based grading provides (Heflebower et al., 2014; Schimmer, 2016). One of the most challenging processes during the standards-based grading implementation phase is educating and garnering the support of the community (Guskey et al., 2014; Marzano, 2000; Spencer, 2012). Another challenge that presents itself during educational reform is the impact that change has on the students (Mitra, 2008).

Academic self-efficacy is the belief in one's own ability to accomplish an academic task (Bandura, 1986, 1993, 1997; Dweck, 1986; Honicke & Broadbent, 2015; Pajares, 1996; Schunk & Pajares, 2002). This belief in oneself has been correlated with academic achievement, student choice in challenging activities, motivation, and perseverance (Bandura, 1986, 1993; Dweck, 1986; Honicke & Broadbent, 2015; Pajares, 1995; Schunk & Pajares, 2002). Achievement levels in past experiences do not necessarily transfer into high self-efficacy in future tasks, thus eliminating the idea that low self-efficacy is a low-achiever problem (Dweck, 1986). Moreover, Dweck (1986) showed that, when students are confused during the first attempt at the completion of a task, the likelihood of the student reaching mastery is seriously threatened. The impact that an individual's self-efficacy has on behavior is observed in a person's choice of activity and is evidenced in their coping mechanisms utilized once they are engaged in that activity (Bandura,

1977). That is, a person will choose activities which they believe they can be successful at, and, when challenged in an activity that they have low self-efficacy in, they will struggle to cope with adversity. Self-efficacy is not the only influence on a person's level of achievement on a specific task; however, if a person has the appropriate skills, then self-efficacy becomes a significant factor in predicting success (Bandura, 1977).

Statement of Problem

The intent of grading is to have a system that communicates student progress to all individuals involved in the academic interests of a student (Brookhart, 2009; Guskey, 2015; Guskey & Munoz, 2015; Kohn, 2011). Specifically, grades are in place to communicate a student's academic progress to parents and others, provide feedback to students for reflection, identify and place students on educational paths, incentivize students, evaluate programs, and document a student's efforts (Guskey, 2015). Grades have become a critical pillar of the American educational system, remaining largely unchanged over the past century, even in spite of educational scholars' recognition of the deficiencies that traditional grades offer (Guskey, 2015). Educational researcher Robert Marzano (2000) notes, despite the inaccuracies that traditional grades present, they are widely accepted by most Americans (Marzano, 2000). He goes on to argue that in fact, "grades are so imprecise that they are almost meaningless" (Marzano, 2000, p.1). Despite the challenges and scrutiny that grades have faced, grading reform has proven to be incredibly difficult because the consequences of poor implementation could have a catastrophic impact on student promotion, college admissions, and entrance into advanced courses (Guskey, 2015).

Teachers report very low confidence levels in grade reporting when asked about accuracy, meaningfulness, and correlation to district or state adopted content standards (Allen,

2005; O'Connor, 2007). Grades have been recognized by statisticians as a good example of unreliable forms of measurement (Brookhart, 1993). Traditional grades become unreliable and invalid due to non-academic factors that influence the grade, individual weighting of assessments, and teacher interpretation of scores on classroom assessments (Marzano, 2000). Establishing a grading process that has clearly defined performance goals and criteria, increases student achievement and provides a foundation for fair and equitable reporting (Guskey & Munoz, 2015; Hattie, 2012; O'Connor, 2007).

In response to much of the criticisms that traditional grades have received, some schools have looked to implementing a standards-based grading model. Standards-based grading is a system of grading and reporting student progress based on their proficiency level on designated content standards (Guskey & Jung 2013; Townsley, 2019). Student grades are reported, generally on a scale from 0-4 for each content standard as opposed to an average of scores on all content standards (Brookhart, 2009; Cox, 2011; Guskey & Jung, 2013; Guskey & Munoz, 2015; Heflebower et al., 2014; O'Connor, 2007; Schimmer, 2016). Standards-based grading systems are based completely on student achievement toward specific standards, thus eliminating behaviors in grading and allowing students to reassess as needed to demonstrate proficiency. Research has shown that teachers believe standards-based report cards provide more beneficial information to parents as compared to traditional report cards (Guskey, 2011). Despite the clarity in communication that standards-based report cards offer, schools are often met with resistance when reforming long held grading practices (Peters & Buckmiller, 2014). The fundamental goal of standards-based grading is to evaluate student achievement impartially, using similar benchmarks for all levels (O'Connor, 2007; Heflebower et al., 2014). Several studies and books have been written examining effective practices for implementing this new

grading system, including direction on guiding parents, teachers, and school leaders through this transition (Cox, 2011; Guskey & Munoz, 2015; Heflebower et al., 2014; O'Connor, 2007; Schimmer, 2016).

The deficiencies associated with the traditional grading system have been recognized by professionals in the field of education for over a century, with very little change taking place (Finkelstein, 1913; Guskey, 2013; Starch & Elliot, 1912, 1913). One of the difficulties with transitioning grading systems today is the deeply entrenched beliefs that have been developed while allowing the traditional grading system to remain (Brookhart, 2009, 2011; Guskey, 2013; Marzano, 2000; Schimmer, 2016). The transition from a traditional grading system to a standards-based system in high schools has been much slower than the same transition within elementary schools (Cox, 2011). This could be because one of the fundamental applications of grades at the high school level is to determine college readiness. As grading reform is addressed, high schools must analyze educational foundations such as honor rolls, eligibility, and academic excellence awards while investigating and implementing grading reform measures (Guskey, 2013; Peters, Kruse, Buckmiller, & Townsley, 2017).

Aside from high school grade point average and Scholastic Aptitude Test (SAT) scores, another significant factor in predicting college success is academic self-efficacy (Chemers, Garcia, & Hu, 2001; Gore, 2006; Hannon, 2014; Honicke & Broadbent, 2015). Albert Bandura formulated the Self-Efficacy Theory in 1977 by analyzing human learning behaviors and motivation as a cerebral function (Bandura, 1977). Self-efficacy is a person's confidence in their abilities to successfully complete a challenging task (Bandura, 1977), whereas academic self-efficacy is described as "the belief in one's capabilities to organize and execute courses of action required to produce given attainments" (Bandura, 1997, p.3). Research has shown that the

construct of self-efficacy is not only closely related to academic achievement but that it also impacts a person's motivation and perseverance towards a task in an academic setting (Bandura, 1986, 1993; Dweck, 1986; Honicke & Broadbent, 2015; Pajares, 1995; Schunk & Pajares, 2002). Bandura (1993) states that self-efficacy ideals impact college success by increasing or decreasing motivation and persistence to complete challenging tasks.

Despite the fact that several studies have shown that academic self-efficacy is a vital factor to consider when speculating on a student's educational potential (Bandura, 1997; Hannon, 2014; Honicke & Broadbent, 2015; Pajares, 1995; Pajares & Usher, 2006;), insufficient research has been completed addressing the impact that transitioning into a standards-based grading system has on the academic self-efficacy of the students. Although a strong argument can be made that the traditional grading system observed in high schools is flawed (Brookhart, 2009; Guskey et al., 2014; Guskey & Munoz, 2015; Heflebower & Marzano, 2011; O'Connor, 2007; Marzano, 2000), the impact that changing grading philosophies has on students is still unknown. This mixed methods study augmented the current research on standards-based grading by examining the impact of a standards-based grading system on the academic self-efficacy of high school students in 11th grade English classes.

Background

The primary purpose of issuing grades to students in K-12 education is to communicate academic progress to students, parents, teachers, and administrators (Brookhart, 2009; Guskey, 2015; O'Connor, 2007). Despite state and federal oversite in many areas of education, schools in every state have the responsibility to design a report card that effectively communicates student academic performance (Swan et al., 2014). The common school movement during the 1800s created the need for teachers to communicate with an increasingly large group of stakeholders

regarding student academic progress, and thus came the advent of the report card (Brookhart, 2009). As a tool to support more effective communication, a grading system was developed for reporting student progress, which quickly became common by the late 1800's, and eventually evolved into the traditional (A, B, C, D, & F) grading system (Brookhart, 2009; Guskey, 1996; Starch & Elliot, 1912, 1913). Despite critical research regarding its validity and reliability, this traditional system of communicating student progress has remained largely unchanged since its inception and has become a pillar of K-12 education in America (Brookhart, 2009; Starch & Elliot, 1912, 1913).

In 1983, the National Commission on Excellence in Education published a report titled *A Nation at Risk.* This report identified the lack of rigorous standards or expectations as a key area of focus to improve K-12 education in the United States (Schimmer, 2016). The impact of the report forced educators and educational institutions to examine curricular standards and expectations in an effort to determine what students should be learning in each curricular area (Schimmer, 2016). In the 1990's, most states had adopted some type of content standards for each subject area at every grade level as a guide for all stakeholders to know what students should be learning and what should be taught by teachers (Schimmer, 2016). These guidelines forced educational institutions to implement standards-based models of instructions with the intent of each class teaching the same content standards (Brookhart, 2009; Heflebower et al., 2014; Marzano, 2000; Schimmer, 2016). Although K-12 education has effectively transitioned instructional practices into a standards-based model, grading has failed to make the same advances, and, thus, teachers are instructing in a standards-based model but failing to report in a manner that supports this method (Brookhart, 2009; Heflebower et al., 2014; Marzano, 2000; Schimmer, 2016).

With a strong argument being made that the traditional grading system observed in high schools is flawed (Brookhart, 2009; Guskey et al., 2014; Guskey & Munoz, 2015; Heflebower & Marzano, 2011; O'Connor, 2007; Marzano, 2000), a system known today as standards-based grading was developed to address some of the deficiencies observed in traditional grading systems. Standards-based grading is a system of reporting grades on a small scale (usually 0-4) for each standard in a specific curricular area (Brookhart, 2009; Heflebower et al., 2014; Marzano, 2000; Schimmer, 2016). A student receives specific feedback on each standard within a course, represented as 0-4 rather than a cumulative grade for the course. These standards-based grades are designed to communicate student progress on each standard within a course. Other characteristics of standards-based grades that are incorporated in this philosophy include the following: removal of behavior attributes from grades, removal of extra-credit, and the elimination of the averaging of scores (Guskey, 2015). A standards-based grading system bases grades purely on a student's ability to reach proficiency at their own pace, without penalty for mistakes through the process, thus eliminating the need to average grades (Heflebower et al., 2014; O'Connor, 2007; Schimmer, 2016). Parents, students, and teachers have been shown to favor the communication provided in a standards-based grading system over a traditional grading system (Swan et al., 2014).

Despite the support for grading reform, schools, districts and states have experienced challenges through the implementation process of standards-based grading (Brookhart et al., 2016; Schimmer, 2016). Although many teachers recognize the need for grading reform, implementation of new grading practices has lacked fidelity (Cox, 2011; Guskey, 2015). Parents are also often opposed to making changes to the grading system, forcing schools to slow down to keep all stakeholders informed and educated throughout the transition process (Franklin et al.,

2016; Peters et al., 2017). Although a strong argument can be made that the traditional grading system observed in high schools is flawed (Brookhart, 2009; Guskey et al., 2014; Guskey & Munoz, 2015; Heflebower & Marzano, 2011; O'Connor, 2007; Marzano, 2000), very little research has been conducted regarding the impact that transitioning from a traditional grading system into a standards-based grading system has on the students involved.

The theory of individual self-efficacy, originally developed by Albert Bandura in 1977, has been linked in studies as an indicator of student academic success (Bandura, Claudio, Caprara, & Pastoreli, 1996; Pajares, 1995; Pajares & Usher, 2006). Academic self-efficacy has been shown to be a characteristic that can be developed and destroyed through a variety of experiences (Bandura, 1986). This study looked at the impact of standards-based grading on a student's academic self-efficacy in English classes during their junior year of high school.

Research Questions

Much information has been reported regarding the inaccuracies of traditional grades and the benefits of transitioning into a standards-based grading system (Brookhart, 2009; Guskey, 2013; Kohn, 2011; Nitko & Brookhart, 2007). Reforming entrenched grading systems has proven to be a challenging transition for both students and parents alike (Cox, 2011; Peters et al., 2017; Schimmer, 2016). Research has shown that standards-based grading systems have benefits, such as the clear communication of student achievement (Brookhart, 2009; Cox, 2011; Guskey & Munoz, 2015; Heflebower et al., 2014; O'Connor, 2007; Schimmer, 2016). It is imperative to understand the impact that such a change may have on student academic success. The intent of this study was to measure the impact that changing grading systems has on the academic self-efficacy of students. Students at two different high schools with similar demographics, were asked to answer a 10 questions General Self-Efficacy instrument in a post

then pre retrospective design. Classes were chosen for the study based on the application of a standards-based grading philosophy in the classes. This method of capturing the impact of standards-based grading on a student's self-efficacy was determined to be appropriate because we are looking at their perceptions, attitudes, and behaviors (Colosi & Dunifon, 2006). At the end of the survey, all students were asked if they would be interested in participating in semi-structured qualitative interviews. A group of students at each school were contacted to participate in the semi-structured interviews to provide qualitative data through student voice.

This mixed methods study looked at the impact that a standards-based grading system has on the academic self-efficacy of high school junior English students. The study examined the following questions:

1) What is the impact of standards-based grading on a high school student's academic self-efficacy?
2) What relationships can be drawn between variables associated with the high school student population of schools in a standards-based grading system compared to their previous experience in a traditional grading system?

Definition of Key Terms

It is both beneficial and vital that educators have a shared common language when addressing academic topics (Creswell, 2015; Marzano & Boogren, 2010). When discussing standards-based grading and issues around grading reform, terms and concepts must be clarified (Brookhart, 2009; Heflebower et al., 2014; O'Connor, 2007). The definitions of terms related to the study provide a guideline and clarity for the reader (Creswell, 2015). Using accurate terminology improves understanding and reduces confusion (Marzano, 2000). Therefore, the

following terms are offered as clarification for the use and meaning of terms within this research study.

Standards-based grading: A system of grading that describes student academic progress in relation to specific standards. These grades are not averaged, but, instead, are specific to a student's achievement level on an identified standard. For the purposes of this paper this term will be used interchangeably with standards referenced grading (Heflebower et al., 2014).

Traditional grading: A system of assigning letter grades (A, B, C, D, and F) to represent student achievement on an assignment or in a class. These grades represent averages and often times allow student behaviors to influence the reporting (Brookhart 2009; O'Connor, Jung, & Reeves, 2018).

Score: A number or letter grade assigned to a student which is factored into the overall grade in the future (O'Connor, 2007).

Grade: A number or letter assigned at the end of a grading period intended to communicate student achievement in a specific subject matter (O'Connor, 2007).

Validity: The degree to which grades convey meaningful and appropriate information regarding a student's academic progress (Brookhart, 2009).

Reliability: The degree to which grades convey accurate information regarding a student's academic progress (Brookhart, 2009).

Self-efficacy: A person's belief in their own ability to systematically complete the steps required to successfully finish a task (Bandura, 1997). Specifically, in the case of this research, in the area of academics.

GPA: Grade Point Average is the average of all grades over a specific period of time (semester, year, high school years).

SAT: Scholastic Aptitude Test is a commonly used standardized exam, widely used as a factor in college admissions.

Significance of the Study

This study is of particular value to schools, school districts, and policy makers at all levels as they look to reform grading practices. A significant amount of research has been done to demonstrate that standards-based grading systems are more effective than traditional grading systems at communicating a student's academic progress (Brookhart, 2009; Cox, 2011; Guskey & Munoz, 2015; Heflebower et al., 2014; O'Connor, 2007; Schimmer, 2016). Although the aforementioned research supports the value in clearly communicating academic progress to all stakeholders, it does not address the impact that this communication has on the academic self-efficacy of students. This study builds upon the existing research by analyzing the impact of grading reform on a high school student's academic self-efficacy before and after the implementation of standards-based grading. The study captures student voice and identifies specific barriers that students face and that educational institutions must address throughout the transition process. School and district leadership will find this study particularly valuable as they look to transition grading systems. It gives foresight to potential issues that must be addressed throughout the process. Policy makers will also find value in this research as they look at all aspects involved in grading reform. While it accepts the need for grading reform, the study looks at this change from a student's perspective, specifically regarding the impact on a student's academic self-efficacy.

Theoretical Framework

Albert Bandura (1977) developed the theory of self-efficacy, defined as a person's belief in their own ability to accomplish a task (Bandura, 1997). A person's level of self-efficacy

impacts their choices in challenging activities, motivation, and reaction to set backs (Bandura, 1997). These beliefs have been shown to be a vital factor in a person's ability to successfully complete a task and their overall competence (Bandura, 1997). The implications of both high and low self-efficacy can have extensive impact in determining how a person addresses challenging situations (Bandura, 1997).

In the academic setting, self-efficacy has been shown to play a major factor in the overall academic success of a student (Bandura et al., 1996; Pajares, 1995; Pajares & Usher, 2006).

> Self-efficacy beliefs are constructed from four principal sources of information: enactive mastery experiences that serve as indicators of capability; vicarious experiences that alter efficacy beliefs through transmission of competencies and comparison with the attainments of others; verbal persuasion and allied types of social influences that one possesses certain capabilities; and physiological and affective states from which people partly judge their capableness, strength, and vulnerability to dysfunction. (Bandura, 1997)

This study looked at the implications of a standards-based grading system on a high school junior student's mastery experiences and overall academic self-efficacy in English classes.

Overview of Research Methods

This study used a mixed methods approach to analyze the impact of a standard-based grading system. More specifically, the study looked at the impact that transitioning into standards-based grading has on the academic self-efficacy of high school juniors. The quantitative data was obtained through the administration of an adapted version of the General Self-Efficacy Survey (GSE) (Appendix A) that was designed to elicit student views on the grades

they receive and the impact that these grades have on their academic self-efficacy. This survey was administered using the retrospective post-pre survey methodology, in which students were asked about their current academic self-efficacy and then were asked to reflect on their academic self-efficacy prior to transitioning into a standards-based grading model. At the end of the survey, students were asked if they would be interested in participating in semi-structured interviews at a future date. Students were then chosen to participate in semi-structured interviews to obtain data for the qualitative portion of the study. The results of both the quantitative and qualitative data obtained were kept independent of each other until the final data analysis to avoid bias during the coding of the qualitative data. In a mixed methods convergent approach, the researcher analyzed both data sets separately and then compared the results to determine if the data converged (Creswell & Guetterman, 2019). This convergent, mixed-methods design was adopted to develop a complete understanding of the research problems by allowing the strengths of the qualitative research to offset the weaknesses of the quantitative research and vice versa (Creswell & Guetterman, 2019). After data collection, convergence or divergence of the data was determined as an instrument to provide the researcher detailed understanding of the proposed research questions.

Chapter II

Review of Literature

Introduction

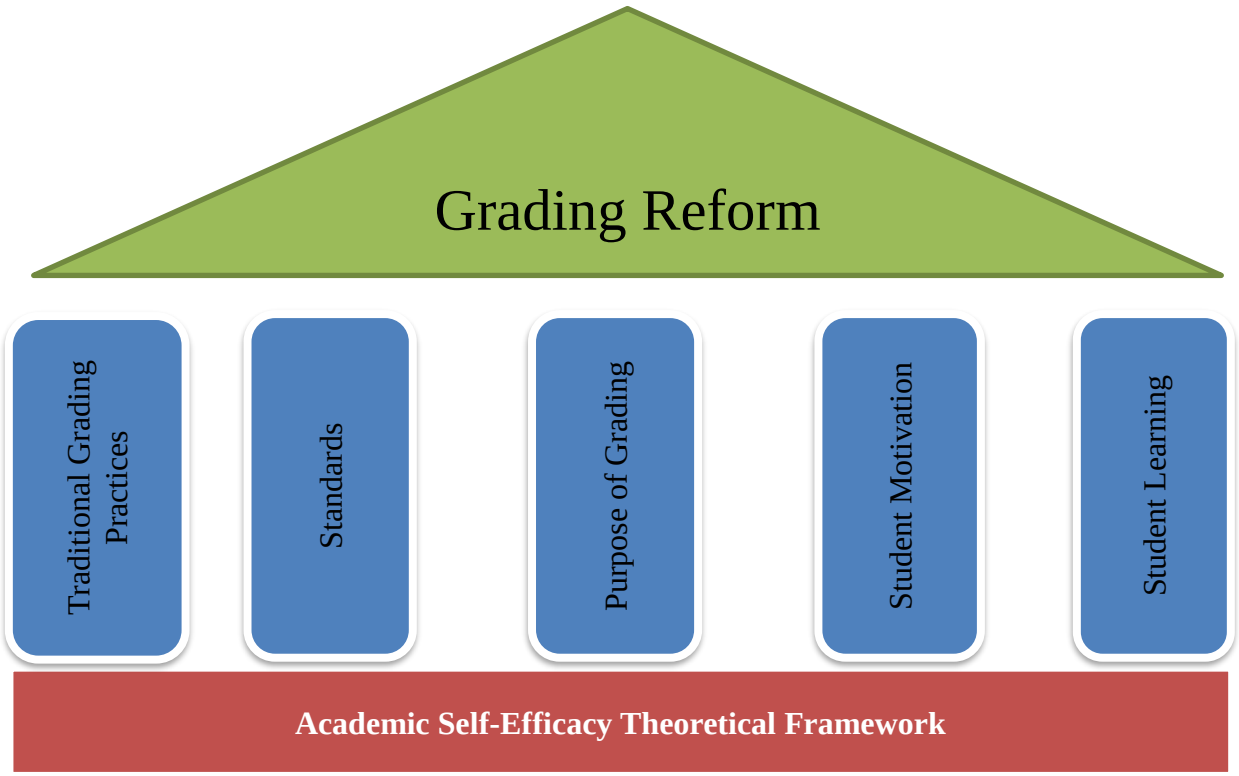
The purpose of this study was to understand the impact that standards-based grading has on the academic self-efficacy of high school students. The addition of specific and rigorous state adopted standards has helped provide clarity for student achievement targets for each grade (Davis, 2019). Reforming grading practices to accurately reflect academic achievement specific to each standard has been studied deeply over the past ten years (D'Agostino et al., 2013; Heflebower & Marzano, 2011; Munoz & Guskey, 2015). These studies have concluded that systems which accurately communicate academic achievement such as standards-based grading lead to increases in student achievement (D'Agostino et al., 2013; Munoz & Guskey, 2015). Though much research has been conducted on the implementation of standards-based grading, very little has looked at the impact that this grading system has on the academic self-efficacy of students (Guskey, 2001; Pollio & Hochbein, 2015; Scarlett, 2018; Scriffny, 2008). Academic self-efficacy has been linked to student success factors such as motivation and perseverance, as well as a student's willingness to take on a difficult task. Analyzing the connection between a specific grading system and a student's academic self-efficacy will provide insight to policy makers and leaders as they look to reform current grading practices.

This literature review will present an overview of traditional grading practices and deficiencies that have led schools throughout the United States to reform classroom grading and adopt standards-based grading philosophies to improve the reporting of student achievement. The implementation and mindset shift that is required to adjust the profoundly ingrained practices of traditional grading poses a challenge for all stakeholders (Franklin et al., 2016;

Peters et al., 2017; Schimmer, 2016; Swan et al., 2014). Although research has clearly evidenced the inadequacies of traditional grading systems, as well as the benefits for all stakeholders that would come with moving towards a standards-based grading model, many educational institutions maintain the use of the traditional grading model (Brookhart, 2009; Marzano, 2010; O'Connor 2007; Schimmer 2016). The failure of schools to transition into a more transparent grading system can be attributed to the many variables involved in creating educational change. This study specifically analyzed the impact that reforming grading practices into a standards-based system has on the academic self-efficacy of high school juniors in an English classroom. The impact that changing grading systems has on students will be analyzed through the following: 1) impact on the student's academic self-efficacy, 2) the student's ability to overcome challenges, and 3) the effect of student perceptions of fairness on academic achievement. This review will provide a comprehensive overview of traditional grading practices and the limitations that experts have identified. It will also provide an understanding of the foundation of current research surrounding the best practices for effective implementation of standards-based grading, including the identification of areas where insufficient research has been conducted, thus validating the need for this study. Figure 2 (shown below) provides a visual overview of the following review of literature.

Figure 2.

Categories of the Literature Review

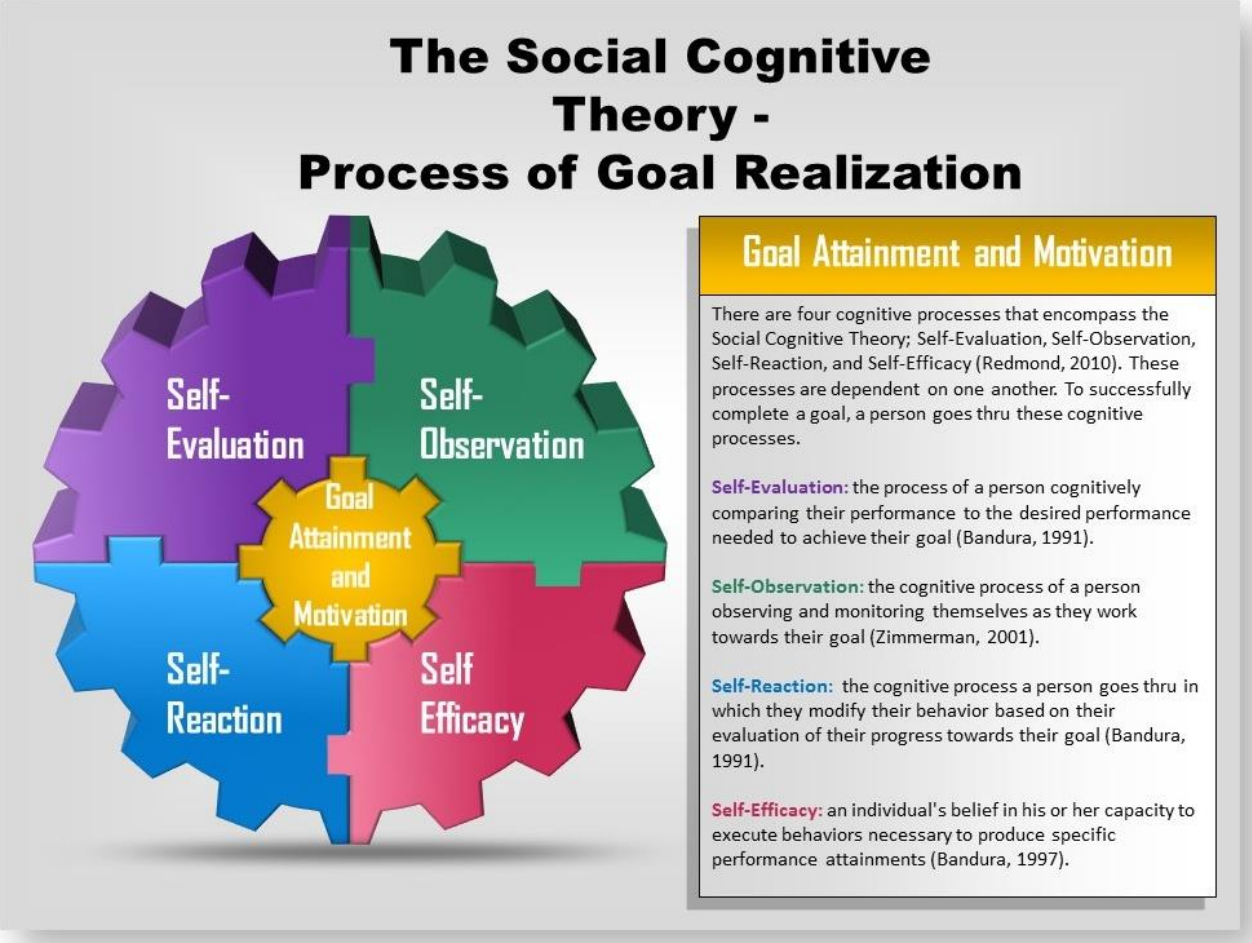


Theoretical Framework

Educational psychologist Albert Bandura, developer of the social learning theory, refers to self-efficacy as a person's belief in their own abilities to successfully accomplish a task (Bandura, 1997). Academic "self-efficacy is defined in terms of individuals' perceived capabilities to attain designated types of performances and achieve specific results" in academics (Pajares, 1996, p. 546). Self-efficacy is one of four major processes which creates the foundation for Social Cognitive Theory as it pertains to goal achievement and motivation (Redmond, 2010). These four processes, as illustrated in Figure 3 below, are self-evaluation, self-observation, self-reaction, and self-efficacy (Redmond, 2010).

Figure 3.

The Social Cognitive Theory- Process of Goal Realization



Note. Adapted from Redmond, B., last modified by McNabb, H. (2010-2016). PSYCH 484 Wiki: 7. Self-Efficacy and Social Cognitive Theories. Retrieved on February 27, 2016 from: 7. Self-Efficacy and Social Cognitive Theories.

Self-efficacy has proven to be a reliable predictor of both motivation and production over an extended period of time (Bandura & Lock, 2003). Self-efficacy has been studied significantly in classroom settings and determined to be a major factor in a student's academic achievement

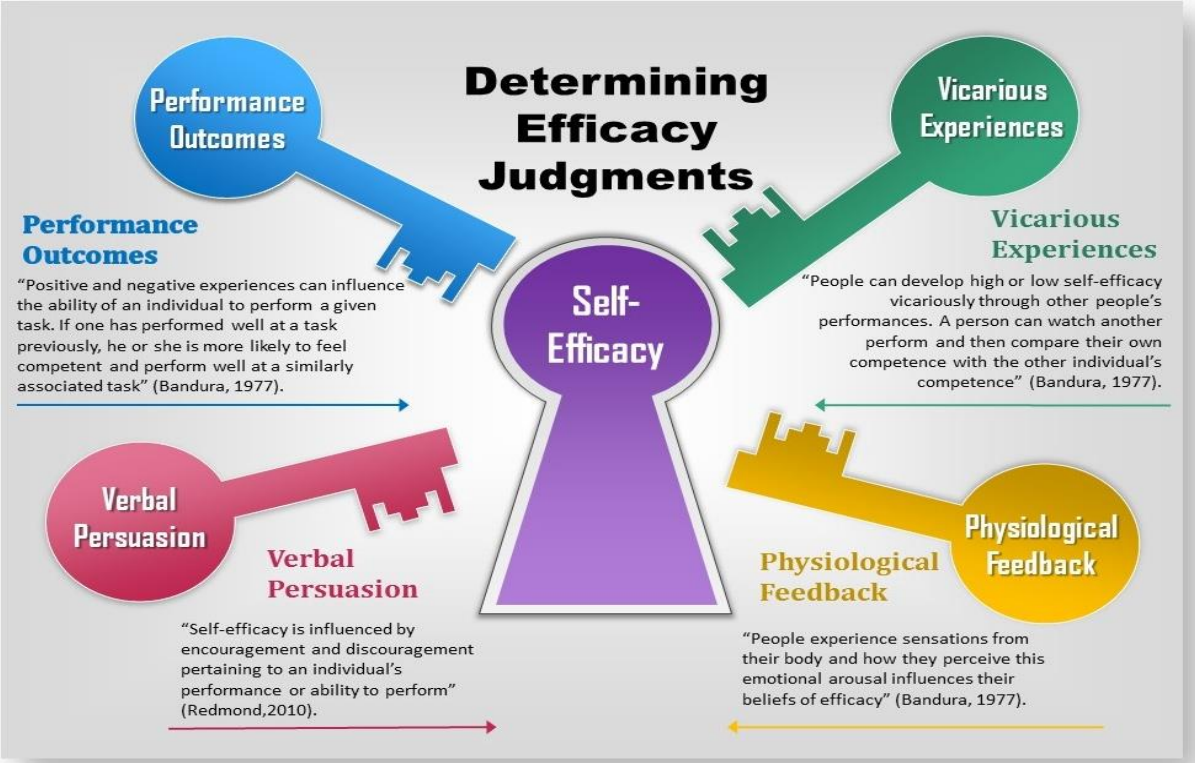
(Bandura & Schunk, 1981; Hannon, 2014; Honicke & Broadbent, 2015; Komarraju & Nadler, 2013; Pajares, 2009; Pajares & Schunk, 2001). In academic settings, the power of social and personality factors, such as self-efficacy, have been correlated with college success (Gore, 2006; Hannon, 2014). Academic self-efficacy has been shown to have a role in predicting the academic achievement of students at all levels, but less is known about how a robust academic self-efficacy is developed and fostered (Pajares & Usher, 2006). These academic self-efficacy beliefs are impressionable and can be impacted both positively and negatively by a person's experiences (Bandura, 1997). One study conducted on the academic self-efficacy of middle school students in the southeastern United States showed that students' perceived mastery in a content area or on a specific standard was an accurate predictor of their effort toward subsequent learning opportunities (Pajares & Usher, 2006). Other studies show that academic self-efficacy can be heightened by allowing students to set specific, obtainable goals (Bandura & Schunk, 1981). Both goal setting and achievement of specific content standards are key components of a standards-based grading system (Heflebower et al., 2014; Gusky & Munoz, 2015; O'Connor, 2007; Schimmer, 2016). The specific impact of standards-based grading on a student's academic confidence and, thus, their effort toward future learning has not yet been studied.

Self-efficacy and self-esteem are distinctly different but are often times confused (Bandura, 1997). Self-efficacy is a person's confidence in their own capabilities, whereas self-esteem is one's perception of their own self-worth (Bandura, 1997). Self-efficacy theory explains that individuals determine their own efficacy based on their prior experiences, performances and other physiological indicators (Schunk, 1991). Success increases a person's efficacy and failures bring it down; however, the impact of failure is decreased if a person has established a high self-efficacy (Bandura, 1986). Figure 4 (shown below) illustrates the four

major components that determine a person's self-efficacy: performance outcomes, vicarious experience, physiological feedback, and verbal persuasion.

Figure 4.

Determining Efficacy Judgements



Note: Adapted from Redmond, B., last modified by McNabb, H. (2010-2016). PSYCH 484 Wiki: 7. Self-Efficacy and Social Cognitive Theories. Retrieved on February 27, 2016 from: 7. Self-Efficacy and Social Cognitive Theories

The effects of self-efficacy on a person's actions can be viewed in a variety of ways. People generally use predictive cognitive measures to influence their capabilities and set personal achievement goals (Bandura, 1991). People with higher self-efficacy set loftier goals, challenge themselves more, and stay more committed to their activities than those with lower self-efficacy (Bandura, 1991, 1993). People with low self-efficacy often tend to avoid difficult

tasks and have less of a commitment toward the goals that they set (Bandura, 1993). Those with low self-efficacy often dwell on their areas of weakness and possible challenges that a task might present rather than concentrating on how to successfully complete the task (Bandura, 1993). The opposite is true for those who have a high level of self-efficacy (Bandura, 1993). People with a high self-efficacy tend to work harder to achieve their goals and are usually able to overcome failure without giving up (Bandura, 1993). Research shows that self-efficacy in children, specifically, can impact academic performance, peer relationships, decision making, and social emotional behaviors (Bandura et al., 1996)

Self-efficacy levels impact a person in three specific ways (Pajares, 1996; Schunk, 1991). First, they impact a person's behavior because people choose to engage in activities that they feel confident in (Bandura, 1991; Pajares, 1996; Schunk, 1991). Second, they impact a person's willingness to complete specific tasks because they determine how much motivation and determination a person has toward a specific activity (Bandura, 1991; Pajares, 1996; Schunk, 1991). Lastly, individuals identified as having low self-efficacy tend to look at problems as being more complicated than they actually are, which may lead to stress and depression (Bandura, 1991; Pajares, 1996; Schunk, 1991). Those with high self-efficacy remain level headed and focused on the task at hand when dealing with challenging work (Bandura, 1991; Pajares, 1996; Schunk, 1991).

Research shows that "children's belief in their academic efficacy is linked to scholastic achievement both directly and through its impact on academic aspirations, prosocial conduct, and lowering proneness to despondency" (Bandura et al., 1996, p.1215). The relationship between academic self-efficacy and academic achievement has been analyzed by many researchers (Chemers et al., 2001). Chemers et al. (2001) found significant evidence that high self-efficacy

positively correlated with academic performance in a sample of first-year university students. This concept that academic self-efficacy is correlated with academic achievement can be used to develop interventions that focus on building a students' academic self-efficacy to impact their academic achievement (Di Guinta et al., 2013).

When comparing students with different levels of self-efficacy, students with high self-efficacy participate more in learning activities, work harder, and show more determination when they are challenged with difficult tasks (Pajares, 1996; Schunk & Pajares, 2002; Schunk, 1991). When evaluating the effects of self-efficacy in academic settings, it is crucial to look specifically at academic self-efficacy, referring to a student's belief in themselves to carry out academic tasks, rather than looking at generalized self-efficacy (Espenshade, Lynch, & Zajacova, 2005). Students who have high academic self-efficacy show better work ethic, monitor their learning progress, and employ self-regulation strategies that encourage academic achievement (Pajares, 2002).

This concept of self-efficacy has been broadly addressed in recent books, such as Carol Dweck's (2006) *Mindset: The New Psychology of Success* and Angela Duckworth's (2016) *Grit: The Power and Passion of Perseverance*, which illustrate the importance of the essential skills of relentless determination and positivity through challenging situations. Carol Dweck (2006) introduced an idea of mindset, and, more precisely, the difference between a growth mindset and a fixed mindset. "This growth mindset is based on the belief that your basic qualities are things you can cultivate through your efforts" (Dweck, 2006, p. 7). Dweck (2006) explains that these mindsets are taught and learned over time and dictate how people react to learning challenges. Throughout her book, Dweck analyzes how this growth or fixed mindset affects people. Students with a fixed mindset were observed wanting to avoid challenging work due to the fear

of failure, while those students with a growth mindset want the challenging work with the understanding that failure will help them grow (Dweck, 2006). She explains that the fixed mindset is about the outcome, whereas the growth mindset is about the process (Dweck, 2006).

Growth and fixed mindsets are observed in sports, education, relationships, and business. People with fixed mindsets want to stay away from difficult challenges due to their overwhelming fear of failure instead of looking at failure as an opportunity to learn. Parents often praise their children for their brains or talents which could create problems when the children face anything tricky or challenging. Dweck encourages parents to focus on the challenges that children face and praise the struggle as an opportunity for learning (Dweck, 2006). The good news, as Dweck (2006) explains, is that the growth mindset can be learned. For the purposes of this literature review, mindset and self-confidence are significantly dependent on each other. Students who embrace challenges as a learning tool have higher self-confidence and academic success in school.

Angela Duckworth (2016) uses the word "grit" to describe the unique quality that makes a person successful or not successful. Duckworth explains that this idea of grit is the combination of passion and persistence towards a goal over a long period of time (Duckworth, 2016). This idea was developed after Duckworth studied a variety of sample groups in high-stress situations. This study was able to predict, with a moderate level of accuracy, which people would or would not be successful based on the results of a grit survey that all participants completed (Duckworth, 2016).

Duckworth separated her study into three essential parts, first defining what grit is, then explaining how a person can learn to be gritty, and finally illustrating how someone can be taught to be gritty (Duckworth, 2016). Duckworth (2016) shows that talent is sometimes the

anchor that holds people back from being gritty. People who are incredibly talented at a specific skill or skills may not learn how to be gritty because they have never faced a challenge in that skill; therefore, talent can be a distraction to obtaining the skill of grit (Duckworth, 2016). The grit mentality can be taught and learned, but it is often missed in educational settings (Duckworth, 2016). A gritty person may have the same level of self-efficacy as a person that is less gritty, but the less gritty person's self-efficacy will be more fragile during challenging times (Duckworth, 2016). Using the concept of self-efficacy as the theoretical framework for this study allows the researcher to analyze the impact of grading systems on a student's academic self-efficacy and, in turn, their future academic success.

History of Traditional Grading Practices

Grading, for the purposes of this study, has the primary purpose of communicating student achievement to all stakeholders, and this communication can come in many forms (Bailey & McTighe, 1996; Brookhart, 2009). The traditional system of reporting grades, A, B, C, D, & F, that is most prevalent in American schools today, has been critiqued for its lack of accuracy and reliability almost since its inception in the early 1900s (Campbell, 1921; Guskey, 2013; Starch & Elliot, 1912). Prior to the traditional grading system, student learning was reported to parents by way of personal conversations with respective guardians (Brookhart, 2009; Brookhart et al., 2016; Guskey, 2013). School attendance laws that in the late 19th and early 20th century dramatically increased the number of students attending school, thus making the reporting of grades through conferences unfeasible (Brookhart, 2009; Brookhart et al., 2016; Gutek, 1986). As a mechanism to communicate student progress to all stakeholders, teachers began using percentages and letter grades to verify and communicate student achievement in each individual subject (Brookhart, 2009; Brookhart et al., 2016; Kirschenbaum, Napier, &

Simon, 1971). This led to the development of one of the foundations of current educational practice: the one-hundred-point grading scale with traditional A-F letter grades. Although this system of assessing and communicating student learning has been critiqued extensively, it has remained largely intact over the past century (Brookhart, 2009; Brookhart et al., 2016).

The reporting of grades for students in K-12 education serves four primary purposes: 1) administrative use, 2) student feedback regarding current progress, 3) student direction for future course selection, and 4) to provide formative feedback for teachers to guide their instructional practices (Airasian, 1997; Brookhart, 2009). All of these applications for grading students involve the communication of current and future learning progress to all stakeholders. Although other fundamental purposes for grading exist, this study will focus on the use of grading as a communication tool for student academic progress. Since grades are a valuable resource for all stakeholders to communicate student achievement, it is the responsibility of educators to ensure grades are meaningful, accurate, fair, and support learning (Guskey, 2013; Guskey & Munoz, 2015; Kohn, 2011; O'Connor 2007; O'Connor et al., 2018).

The traditional grading system has been challenged by educational experts for its validity and reporting accuracy almost since its inception over one-hundred years ago (Campbell, 1921;; Guskey, 2013; Kohn, 2011; Starch & Elliot, 1913). Grading has become such a complex issue that some educational experts have argued that grades should be abandoned all together (Kohn, 1999). Critics of the traditional grading system recognize several flaws in the way grades are determined using a one-hundred-point scale. If traditional grades do not accurately communicate student progress, then they are not a true reflection of a student's academic performance (Allen, 2005). Campbell (1921) explains that the marking system in use in 1921 was filled with numerous deficiencies and inaccuracies. This traditional grading system, as described by

Campbell (1921), is mostly intact and in use at a majority of high schools in the United States today (Anderson, 2018). A new cycle of critiques of the traditional grading system came in the 1960s and 1970s, a period of time known for student protest and uprising (Brookhart, 2009). One of the most impactful studies during this time was a book titled *Wad-ja-get? The Grading Game in American Education*, authored by Kirschenbaum et al. (1971). The authors of this scholarly study presented their findings in the form of a novel based around a fictitious school (Kirschenbaum et al., 1971). The novel introduced a new form of grading students and communicating academic performance to stakeholders that is closely aligned with current standards-based grading practices (Brookhart, 2009; Kirschenbaum et al., 1971). During the 1990s and 2000s, federal legislation No Child Left Behind (NCLB) and, more recently, the Every Student Succeeds Act (ESSA) were introduced, requiring states to provide content standards and to report academic progress towards proficiency of these standards (Brookhart, 2009). Although content standards were introduced and put in use at most K-12 schools, traditional grading practices remained the same (Brookhart, 2009). Most recently, grading reform has been focused on the introduction and development of the standards-based grading report card in which student academic progress is communicated for each specific grade level on identified priority standards (Brookhart, 2009; Brookhart et al., 2016; Guskey, 2013).

Deficiencies of Traditional Grading Practices

The ability for a grade to accurately represent and meaningfully communicate a student's achievement toward a specific learning target is referred to as grading reliability and validity (Brookhart, 2009). The reliability of grades from teacher to teacher has been shown to be inconsistent because teachers develop their own grading methods based on personal philosophies (Frisbie & Waltman, 1992). Ken O'Connor (2007) identified fifteen major factors, organized

into four categories that compromise the validity of grade reporting in the traditional grading system. The four practices identified as roots for grade inaccuracies are distorted achievement reporting, low-quality evidence, inappropriate calculations, and learning inhibitors (O'Connor, 2007). Individually, experts have looked at the impact of the various non-academic items that get included in a grade, the misrepresentation caused by giving a student a zero, and the inaccuracies created when a teacher needs to rate a student on an overall one-hundred-point scale. These experts have determined that grade reporting is an unreliable and inaccurate mechanism to report student achievement (Guskey, 2013; Marzano, 2000; O'Connor, 2007; Schimmer, 2016).

In the traditional grading system, teachers average scores from assignments, classwork, homework, quizzes, tests, and projects to determine a letter grade that indicates overall achievement in a particular class (Guskey & Munoz, 2015; Kunnath, 2017). Individual grades are compiled over a period of time, and a final grade is issued by combining all of these factors into one final grade (Kunnath, 2017). Due to the autonomous but sophisticated methods behind grading, the grade that the student receives in a specific class often leaves educators, parents, and students uninformed about a student's progress (Kunnath, 2017). It also may indicate that a student is passing even if they lack the foundational skills necessary to be successful in future courses (Kunnath, 2017). Teacher-student conversations tend to revolve around the point value of an assignment or the possibility of extra-credit rather than a focus on student learning (Scarlett, 2018).

The distorted impact that a zero has on a student's grade in a traditional grading system has been researched and debated heavily over the years (Peters, 2009; Richmond, 2008). School districts have created minimum grading policies, prohibiting a teacher from giving grades lower than 50% (Guskey, 2013). This is recognition of the impact that a single zero can have in the

traditional one-hundred-point scale system. A student receiving a zero in a traditional grading system will need to get perfect scores on nine other assignments to recover (Guskey, 2013). Additionally, zeros are often found in grade books when a student does not complete or turn in an assignment (O'Connor, 2007). If grades are intended to communicate a student's academic performance, a zero is an ineffective way to suggest that the student does not know the concept being assessed; it may be that they simply failed to complete the task (Bailey & Guskey, 2010; O'Connor, 2007). Guskey and Bailey (2010) identify the use of "I" as a symbol for incomplete as a more effective mechanism to communicate student performance. The use of zeros in the traditional grading system does not accurately reflect student performance due to the mathematical inconsistencies of extreme scores (O'Connor, 2007).

The key components of a grade can also vary significantly from teacher to teacher and school to school, leaving parents confused and misinformed regarding their child's achievement level (Anderson, 2018). Due to the lack of specificity in traditional grading, parents are open to their own interpretations regarding the meaning of a grade (Frisbie & Waltman, 1994). In one study, parents identified the average grade as a C+, while the teachers identified a B; explaining that no Ds or Fs were given (Frisbie & Waltman, 1994). This misrepresentation could lead a parent of a C+ student to believe that their child is average when the reality is they are below average in the class (Frisbie & Waltman , 1994). One of the causes of this inconsistency in grading is the use of norm-referenced grading, where students are compared to other students versus criterion-referenced grading where student grades are determined by their ability to show proficiency on a specific standard (Anderson, 2018). Nearly every state in the United States has adopted some form of content standards, but grades are still widely reported by subject area (O'Connor, 2007). Criterion-referenced grading is focused on communicating to all

stakeholders, specific progress on identified goals for each student (D'Agostino et al., 2013; Scarlett, 2018). Critics of traditional grading practices have provided substantial evidence to question the validity, accuracy, and relevance of these grading practices. Grades that are issued today in a traditional grading system should be viewed as almost irrelevant due to their lack of precision (Marzano, 2000).

Rationale for Standards-Based Grading

Standards-based grading was developed in response to the discrepancies identified through the research of traditional grading practices with the intent of creating a system that reports and communicates student progress on specific academic standards (Gusky & Munoz, 2015; Heflebower et al., 2014; O'Connor, 2007; Schimmer, 2016). Recognizing the structural deficiencies that are found in the traditional grading system has helped to create the foundation for development of the standards-based grading model (Gusky & Munoz, 2015; O'Connor, 2007; Schimmer, 2016). The goal of the standards-based grading model is to address the shortcomings of traditional grading by focusing on the reporting of specific skills or standards that are aligned with what students should know and be able to do (Heflebower et al., 2014). This type of academic performance reporting is intended to provide more clarity and specificity to all stakeholders regarding the academic performance of a specific student's progress in an identified standard (Brookhart, 2009).

Standards-based grading is a system that identifies student performance on a specific standard as the single factor that determines the grade, rather than the accumulation and average of several scores over a period of time (Bailey & Guskey, 2010; Brookhart, 2009). In a standards-based grading system, high impact standards, or priority standards, are identified in each curricular area (Heflebower et al., 2014). Students receive a score on a scale from one to

four, which identifies the student's level of proficiency, according to each priority standard (Heflebower et al., 2014). Student scores are then reviewed by the teacher at the end of the grading period to determine an overall achievement score for a particular subject (Heflebower et al., 2014). Unlike traditional grading systems, scores are not averaged; instead, the teacher looks at the preponderance of evidence over a period of time, evaluating recent scores as a more useful indicator of student performance as compared to older scores (Heflebower et al., 2014). This standards-based method of grading students eliminates inequities found in traditional grading by limiting the effects of a zero and doing away with averaging and reporting student progress on specific standards rather than vast subject areas (Heflebower et al., 2014; Schimmer, 2016; O'Connor, 2007). Three critical components of standards-based grading are as follows:

1) Student grades are reported according to their performance on each specific learning standard as compared to an average of all scores to calculate a grade (Beatty, 2013; Knight & Cooper, 2019; Marzano & Heflebower, 2011; Townsley, 2019).
2) Student grades are based on only academic factors. All nonacademic behaviors and effort-based grading is eliminated (Knight & Cooper, 2019; Townsley, 2019).
3) Students are able to make multiple attempts towards proficiency of a standard, without any penalty (Knight & Cooper, 2019; Townsley, 2019)

The foundation of the standards-based grading system must be introduced to stakeholders by what Tom Schimmer (2016) refers to as the standards-based mindset, which not only changes what we grade, but also changes how we grade. The standards-based mindset involves three key elements which include: not averaging grades, eliminating behavior from grading, and making homework meaningful (Schimmer, 2016). Elimination of grade averaging gives a more accurate description of student achievement at a specific moment in time because the grade is dictated by

a student's most recent performance regarding a specific standard (Schimmer, 2016). This transformation also reduces the impact of outlier grades such as zeros that, in a traditional system, have detrimental effects on an overall grade (Schimmer, 2016). Removing student behavior, such as attendance and timeliness, from grading eliminates grade distortion and promotes an uncontaminated assessment of student learning (Iamarino, 2014; Schimmer, 2016). Making homework meaningful, the last pillar of the standards-based mindset, is focused on teachers assigning homework and giving useful, individualized feedback upon completion of each assignment (Schimmer, 2016). Individuals that reach the mastery level of a challenging task have shown the determination of practicing the task repeatedly over time, without the fear of failure (Gladwell, 2008). Homework must be viewed as a formative assignment that allows for mistakes so that students and teachers can obtain valuable feedback to guide their learning and instructional focus (Schimmer, 2016).

The standards-based scoring approach allows teachers to give highly individualized reporting on student achievement based on a product, process, and progress (Guskey & Munoz, 2015). The use of a four-point scale for this reporting, with an easily understood rubric or proficiency scale for each point level that explicitly illustrates indicators at each level, has been identified as a useful tool to support this system (Heflebower et al., 2014; Schimmer, 2016). In general, this four-point scale indicates for each identified standard: 4- exceeds standard, 3- at standard, 2- approaching standard, and 1- not at standard (Heflebower et al., 2014). This method of reporting student learning promotes clear and specific lines of communication about student achievement (D'Agostino et al., 2013). The table below compares and contrasts the key differences in standards-based versus traditional grading systems.

Table 1.

Comparison of Traditional Grading and Standards-Based Grading

Traditional Grading	Standards-Based Grading
1. Based on assessment methods (quizzes, tests, homework, projects, etc.). One grade/entry is given per assessment.	1. Based on learning goals and performance standards. One grade/entry is given per learning goal.
2. Assessments are based on a percentage system. Criteria for success may be unclear.	2. Standards are criterion or proficiency-based. Criteria and targets are made available to students ahead of time.
3. Use an uncertain mix of assessment, achievement, effort, and behavior to determine the final grade. May use late penalties and extra credit.	3. Measures achievement only OR separates achievement from effort/behavior. No penalties or extra credit is given.
4. Everything goes in the grade book, regardless of purpose.	4. Selected assessments (tests, quizzes, projects, etc.) are used for grading purposes.
5. Include every score, regardless of when it was collected. Assessments record the average - not the best - work.	5. Emphasize the most recent evidence of learning when grading

Note: Adapted from O'Connor, K. (2002). How to grade for learning: Linking grades to standards (2nd ed.). Thousand Oaks, CA: Corwin.

Although standards-based grading practices can differ depending on the core philosophies of the educational institution the core's foundation listed above remain the same. Some schools have adopted hybrid systems to allow teachers to give grades that are aligned to the standards-based on evidence obtained over the school year but providing letter grades as a final mark in the class (Heflebower et al., 2014). This is referred to as standards-referenced grading and can be an easier transition and effective implementation.

Implementation of Standards-Based Grading

Reforming grading practices is one of the most effective tools used to improve student achievement due to the impact on both student learning and teacher pedagogy (Guskey & Bailey, 2001; Heflebower & Hoegh, 2014). Despite the well-documented inaccuracies and limitations of the traditional grading system, effective implementation of a standards-based grading system has been challenging for many school districts (Franklin, Buckmiller, & Kruse, 2016; Guskey et al., 2014; Lee et al., 2017). Educators seeking to transform the grading procedures in their schools or districts are working to reshape traditions that are deeply ingrained into the culture of K-12 education (Guskey, 2011). Additionally, educators, students, and guardians need time to adjust to the new grading system (Lee et al., 2017). Guskey (2011) identifies five significant obstacles that must be overcome to implement a standards-based grading system effectively. These five commonly held beliefs are: 1) grades are a tool to separate students, 2) grades should fit a bell curve, 3) grades should provide a comparison amongst other students in the class, 4) bad grades motivate students, and 5) students are given only one grade per course (Guskey, 2011). Districts attempting to implement the standards-based grading model require foundational changes to current processes, habits, and convictions (Heflebower et al., 2014). Change of this magnitude must be handled collaboratively within a school or district, receiving buy-in from leaders, teachers, parents, and students (Heflebower et al., 2014). Moving an organization through this monumental change requires a committed leader that is focused on creating a system that communicates student progress to all stakeholders (Heflebower et al., 2014). Heflebower et al. (2014) illustrate the implementation process in a four-year plan, as seen in Table 2.

Table 2.

Sample Four-Year Standards-Based Grading Implementation Plan

Year One: Curriculum and Communication	Year Two: Capacity Building	Year Three: Implementation	Year Four: Continuation
Identify prioritized standards. Create (or revise) proficiency scales. Create (or revise) quality classroom assessments. Develop a communication plan.	Assemble a guiding team. Uncover current beliefs and attitudes about grading. Establish a group of "scouts" to explore the changes being made and report back Enlist consultants. Educate the board of education.	Announce implementation. Implement new report cards. Encourage small-group experimentation. Organize book studies. Conduct school visits. Establish core beliefs. Involve parents. Involve technology staff.	Implement new teacher development.

Note. Adapted from Heflebower et al., *(2014) A school leaders guide to standards-based grading*. Bloomington, IN: Marzano Research Laboratory.

As illustrated in the table above, the process for any educational institution to transition grading practices must be carefully mapped out with informational opportunities for parents, teachers, school leaders, and board members (Heflebower et al., 2014). The goal of all these informational opportunities is to educate stakeholders regarding the deficiencies of traditional grading practices and to inform them about the benefits of standards-based grading. Schimmer

(2016) refers to this complete overhaul of the way stakeholders view grading students as the "standards-based mindset" shift. The grading paradigm transfers from something that educators sporadically hand out to deliberate communication of a student's learning level on a specific standard at a given time (Schimmer, 2016). This mindset shift is the groundwork that must be well defined and ingrained prior to attempting to change any grading or assessment systems.

Standards-Based Grading for Parents

One of the most difficult obstacles to overcome during the implementation of a standards-based grading system is confusion and the misunderstanding of the standards-based grading philosophy (Guskey & Jung, 2013; Marzano, 2000; Spencer, 2012). Both Tom Schimmer (2016) and Tammy Heflebower et al. (2014) emphasize the importance of getting parental support throughout the grading reform process. Maintaining strong lines of communication with all stakeholders through the transition from traditional grades to a standards-based system is a vital piece of the planning (Schimmer, 2016). Due to a lack of strong communication, parents have resisted grading reform (Spencer, 2012).

Policy makers and educational leaders are recognizing this resistance and starting to include parents in the reform process (Guskey et al., 2014). In Kentucky, before full implementation of standards-based report cards, a sample of parents were selected to participate in a study in which they received both standards-based report cards and traditional report cards (Guskey et al., 2014). Parents were asked to provide feedback regarding both report cards (Guskey et al., 2014). The study analyzed parent perceptions of standards-based grading and found that a majority of parents support this form of grading system. Furthermore, the parents that were not in support of standards-based grading reported strong associations with their in-depth familiarity with traditional grades (Guskey et al., 2014).

Buckmiller et al. (2016) conducted a qualitative study that was able to identify five common themes through a triangulation of data that included researcher notes, participant interviews, and district policy reviews. The first two themes were closely connected; parents expressed comfort in the fact that they have a good understanding of the traditional grading system and extreme discomfort in their unfamiliarity with the new standards-based system (Buckmiller et al., 2016). The third theme was the parental perception of poor communication by the school, which in turn led to a disconnect with stakeholders (Buckmiller et al., 2016). This lack of communication led to the fourth theme, which was confusion among parents and a lack of clarity surrounding the new grading system (Buckmiller et al., 2016). The fifth theme was the perception that students were not academically pushed due to the new retake policy that was a pillar in the new grading philosophy (Buckmiller et al., 2016). Parents noted that they often became frustrated and began to express concerns regarding their child's future, commenting that the standards-based grades did not seem real, mainly because they differed from traditional grades (Buckmiller et al., 2016). The study concluded that the effective implementation of a standards-based grading system is an extremely difficult task for a district or school to undertake (Buckmiller et al., 2016). As institutions look toward grading reform, there must be a plan to include parents in conversations regarding the purpose of grades and grading philosophies prior to making any changes (Buckmiller et al., 2016).

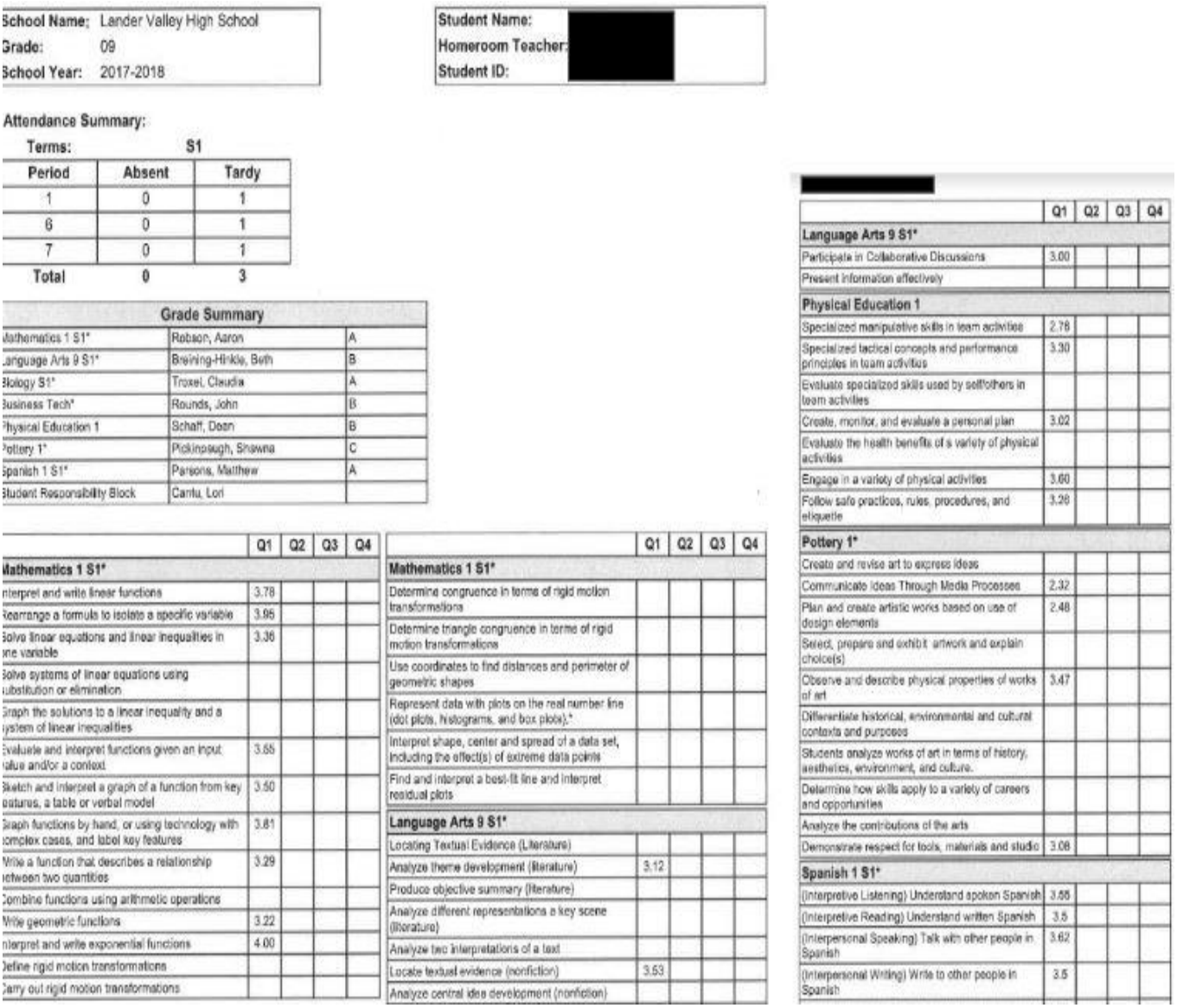
The above-mentioned studies illustrate the need to involve and educate parents regarding grading reform throughout the implementation process. Schools should be intentional about getting the support of all stakeholders, educating them about the shortfalls of traditional grading, discussing the purpose of grading, as well as discussing the benefits of standards-based grading (Schimmer, 2016; Buckmiller et al., 2016). Schools should use all forms of communication to

inform their parents about standards-based grading including websites, forums, emails, and social media (Heflebower et al., 2014). Teachers should also be able explain the new grading system with a high level of proficiency because parents will often begin by seeking more information from the teacher as their first point of contact (Heflebower et al., 2014). It is appropriate to communicate information in parent-friendly language that is easily understood during the implementation process rather than trying to explain information or justify the transition after it has been implemented (Schimmer, 2016).

The following exemplar report cards illustrate what a standards-based grading report card communicates versus what a traditional grading report car communicates. Specific skills such as the students "ability to interpret and write linear functions" are addressed in the standards-based report card and then the skills are factored into determining an overall letter grade. The traditional report card only gives an overall grade, leaving a gap in communication regarding the student's overall strengths and areas of growth specific to each learning target within a course. Figure 5 provides an example of a standard-based grading report card. Figure 6 is an example of a traditional report card.

Figure 5.

Example: Standards-Based Grading Report Card



School Name: Lander Valley High School
Grade: 09
School Year: 2017-2018

Student Name: [redacted]
Homeroom Teacher:
Student ID:

Attendance Summary:

Terms: S1

Period	Absent	Tardy
1	0	1
6	0	1
7	0	1
Total	0	3

Grade Summary		
Mathematics 1 S1*	Robson, Aaron	A
Language Arts 9 S1*	Breining-Hinkle, Beth	B
Biology S1*	Troxel, Claudia	A
Business Tech*	Rounds, John	B
Physical Education 1	Schaff, Dean	B
Pottery 1*	Pickinpaugh, Shawna	C
Spanish 1 S1*	Parsons, Matthew	A
Student Responsibility Block	Cantu, Lori	

	Q1	Q2	Q3	Q4
Mathematics 1 S1*				
Interpret and write linear functions	3.78			
Rearrange a formula to isolate a specific variable	3.95			
Solve linear equations and linear inequalities in one variable	3.36			
Solve systems of linear equations using substitution or elimination				
Graph the solutions to a linear inequality and a system of linear inequalities				
Evaluate and interpret functions given an input value and/or a context	3.65			
Sketch and interpret a graph of a function from key features, a table or verbal model	3.50			
Graph functions by hand, or using technology with complex cases, and label key features	3.81			
Write a function that describes a relationship between two quantities	3.29			
Combine functions using arithmetic operations				
Write geometric functions	3.22			
Interpret and write exponential functions	4.00			
Define rigid motion transformations				
Carry out rigid motion transformations				

	Q1	Q2	Q3	Q4
Mathematics 1 S1*				
Determine congruence in terms of rigid motion transformations				
Determine triangle congruence in terms of rigid motion transformations				
Use coordinates to find distances and perimeter of geometric shapes				
Represent data with plots on the real number line (dot plots, histograms, and box plots).*				
Interpret shape, center and spread of a data set, including the effect(s) of extreme data points				
Find and interpret a best-fit line and interpret residual plots				
Language Arts 9 S1*				
Locating Textual Evidence (Literature)				
Analyze theme development (literature)	3.12			
Produce objective summary (literature)				
Analyze different representations a key scene (literature)				
Analyze two interpretations of a text				
Locate textual evidence (nonfiction)	3.53			
Analyze central idea development (nonfiction)				

[redacted]

	Q1	Q2	Q3	Q4
Language Arts 9 S1*				
Participate in Collaborative Discussions	3.00			
Present information effectively				
Physical Education 1				
Specialized manipulative skills in team activities	2.78			
Specialized tactical concepts and performance principles in team activities	3.30			
Evaluate specialized skills used by self/others in team activities				
Create, monitor, and evaluate a personal plan	3.02			
Evaluate the health benefits of a variety of physical activities				
Engage in a variety of physical activities	3.60			
Follow safe practices, rules, procedures, and etiquette	3.26			
Pottery 1*				
Create and revise art to express ideas				
Communicate Ideas Through Media Processes	2.32			
Plan and create artistic works based on use of design elements	2.48			
Select, prepare and exhibit artwork and explain choice(s)				
Observe and describe physical properties of works of art	3.47			
Differentiate historical, environmental and cultural contexts and purposes				
Students analyze works of art in terms of history, aesthetics, environment, and culture.				
Determine how skills apply to a variety of careers and opportunities				
Analyze the contributions of the arts				
Demonstrate respect for tools, materials and studio	3.08			
Spanish 1 S1*				
(Interpretive Listening) Understand spoken Spanish	3.55			
(Interpretive Reading) Understand written Spanish	3.5			
(Interpersonal Speaking) Talk with other people in Spanish	3.62			
(Interpersonal Writing) Write to other people in Spanish	3.5			

Note. Reprinted from Lander Valley High School report card (2017). Lander, Wy. Retrieved from https://www.allthingsplc.info/evidence/details/id,1418

Figure 6.

Brunswick Academy Report Card

Brunswick Academy

Student: **Felicity Bradshaw**
Class: **9th Graders**
Homeroom Teacher: **Anna Nagy**
Academic Year: **2013-2014**

Report Card

Subject	Cycle 1	Cycle 2	Cycle 3	Exam		Sem 1	Cycle 4	Cycle 5	Cycle 6	Exam		Sem 2	Final	Credit
Algebra II - Joey Little	A	A-		-	-					-	-			-
Fine Arts - Art - Anna Nagy	A	A-		-	-					-	-			-
Honors World Literature - Anna Nagy	A-	A		-	-					-	-			-
Latin II - Joey Little	B+	A+		-	-					-	-			-
Physical Science - Anna Nagy	A-	A-		-	-					-	-			-
World Studies - Joey Little	A	A		-	-					-	-			-

Attendance	Cycle 1	Cycle 2	Cycle 3	Cycle 4	Cycle 5	Cycle 6	Total
Absent (Including Excused)	2	1					3
Tardy (Including Excused)	0	0					0

Grading Scale

97 - 100 = A+	94 - 96 = A	89.5 - 93 = A-	87 - 89.4 = B+	84 - 86 = B	79.5 - 83 = B-
77 - 79.4 = C+	74 - 76 = C	69.5 - 73 = C-	67 - 69.4 = D+	64 - 66 = D	60 - 63 = D-
0 - 60 = F					

Parent's Signature: _____

Note: Brunswick Academy report card (2014). Lawerenceville, Va. Retrieved from https://blog.quickschools.com/2013/12/18/the-brunswick-academy-report-cards/

Standards-Based Grading for Students

There has been an extremely limited amount of research regarding the impact of standards-based grading on students. Research does support the philosophy that grades are an important factor as our students compete internationally on standardized tests and other performance indicators (Vatterott, 2015). With the United States consistently underachieving on standardized tests, when compared to other industrialized countries, and an increasing college dropout rate, it is evident that something is misaligned in our current system (Vatterott, 2015). Students can be valuable contributors to schools that are looking to reform current practices as

they are able inform stakeholders and point out possible deficiencies that may be present (Fielding, 2001; Rudduck & Flutter, 2000). In the most basic form, student voice should be used to inform educational leaders regarding school reform and current educational issues in a collaborative workspace (Mitra, 2008). As this is applied to the transition to standards-based grading, it is imperative to engage in rich dialogue with students in an effort to understand and react to the challenges that are presented.

Mitra (2008) emphasizes the importance of giving students a voice and purposeful input throughout any school reform process. "If the intent is to improve student learning, it defies logic to ignore those most directly affected by grading and assessment decisions; yet students are seldom involved in such discussions" (Peters et al., 2017, p. 12). Peters et al. (2017) found that students were most concerned with five key areas: 1) the implementation process, 2) grading issues, 3) post-secondary preparation 4) social issues, and 5) issues related to teacher pedagogy. Students expressed concerns about the challenges of getting an A, several students felt discouraged because homework was not included in the final grade, and others suggested changing the 4-point grading scale to something different (Peters et al., 2017). Many students also indicated feeling less motivated to do work and felt that the standards-based grading system, in its entirety, is so different from college that it fails to prepare them for post-secondary education (Peters et al., 2017). Student voice is intended to limit the negative perceptions that are presented during educational reform. Studies have indicated that student perceptions of fairness are directly correlated with their learning (Chory-Asad, 2002). Students that felt their teachers were using fair practices in the classroom were more interested in the course (Chory-Asad, 2002). During a transition to a standards-based grading philosophy, it is critical to listen to student concerns and address them in a systematic way to avoid and perceptions or

misconceptions of being treated unfairly. Although students are most directly impacted by substantial transitions in grading systems, very few studies exist that qualitatively give students a voice or quantitatively measure their perceptions to these changes.

Standards-Based Grading and Student Achievement

Although standards-based grades and state assessment scores can provide more accurate communication regarding a student's academic progress, the connection between standards-based grades and assessment scores has not been thoroughly analyzed in academic literature (D'Agostino et al., 2013; Guskey et al., 2014; Townsley, 2019). Studies looking at the impact of standards-based grades on standardized test scores have shown mixed results (Townsley, 2019). The correlation has been observed most prevalently when students scored in the "Meets Standards" category on the state standardized test, indicating that the teachers were able to identify students that met standards according to the standards-based grades that they reported (D'Agostino et al., 2013). Students that fell into the "Falls Far Below Standard" category had the most significant discrepancy between test scores and grades, with mathematics grades having the most significant disconnection (D'Agostino et al., 2013). The study concludes that there may be several reasons for the loose correlation between standards-based grades and standardized test scores, including, but not limited to, the clarity of learning levels for all stakeholders (D'Agostino et al., 2013).

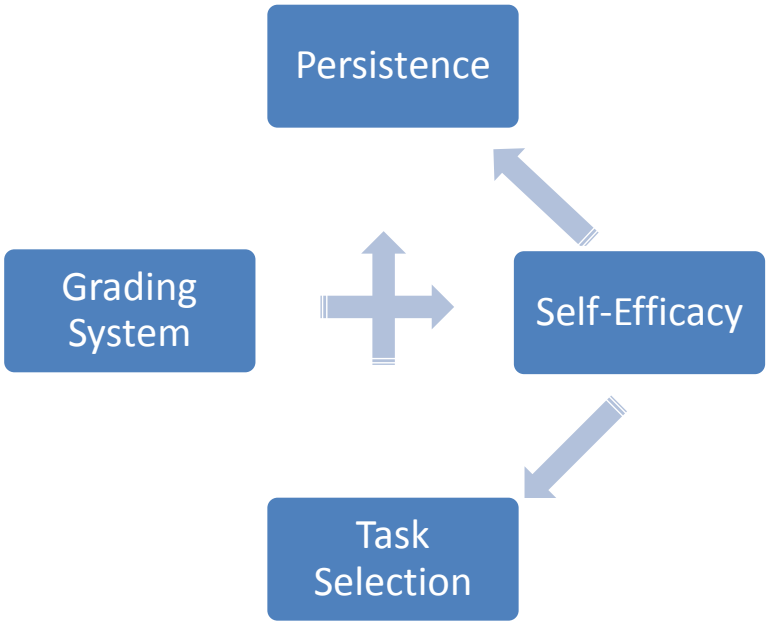
Additional research was conducted comparing the ACT scores and GPAs of two demographically similar high schools in the Midwest, United States (Townsley & Varga, 2018). It found that students in a standards-based system scored lower on college entrance exams as compared to their similar peers (Townsley & Varga, 2018). One possible reason for this finding

was the effect of creating a culture where students can retake exams without exposing students to a high-stakes testing environment (Townsley & Varga, 2018).

Standards-based grading has been associated with standardized testing increases as well. In a large study of schools in the Denver, Colorado area, data showed a stronger correlation between test scores and grades in a standards-based system when compared to similar schools with traditional grades (Scarlett, 2018). Figure 7 shows the interconnected relationship between a grading system, self-efficacy, persistence and task selection.

Figure 7.

Interconnected Relationship Between a Grading System, Self-Efficacy, Persistence, and Task Selection



Student Perceptions of Fairness and Academic Achievement

Extensive evidence exists supporting the notion that student perceptions of grading fairness impact their educational experiences, although very little is known about the teaching and grading practices that are likely to be viewed by students as unfair (Fay & Gordon, 2010). Student achievement in both reading and mathematics has been correlated with their perceptions

of fairness (Ripski & Gregory, 2009). Students' perceptions of hostility have been evidenced to be a valid predictor of scores on standardized tests, in both math and reading (Ripsky & Gregory, 2009). Chory-Assad (2002) concluded that student perceptions of fairness in a classroom setting are a predictor of academic achievement and motivation. Instructors who clearly communicated their fair policies and grading scales had more active student participation in their course (Chory-Assad, 2002). Additionally, evidence supporting the link between perceived classroom fairness and interpersonal aggression was noted (Chory-Assad, 2002). Classroom management and teacher communication style have been shown to closely correlate to effective learning (Chory & McCroskey, 1999). Additional correlations were discovered between non-verbal immediacy and teacher management communication style; and correlations were also found between student attendance and effective learning (Chory & McCroskey, 1999). "These findings suggest that teachers should increase their non-verbal immediacy, delegate decisions to their students, encourage communication between themselves and their students, and urge students to attend class more often" (Chory & McCroskey, 1999, p. 10).

Conclusion

Grades are an essential tool used to communicate academic achievement to all stakeholders. Grades are used as evidence to place students in specific programs, inform stakeholders regarding academic progress, and promote students. Unfortunately, in traditional grading systems, outside factors such as behavior, extra-credit, and the use of zeros tend to distort the information that a grade is intended to communicate (Wormeli, 2011). A student receiving a "B" in math does not effectively communicate what the student knows or does not know. For more than a century, education experts have questioned the traditional grading system commonly used throughout the United States for its validity and accuracy in measuring student

progress. Despite challenges to the current grading system, the way students are graded has been left largely unchanged over the years. In recent years, since the adoption of the No Child Left Behind Act in 2001, education has shifted towards accountability through standardized testing. This testing is centered around standards that the NCLB legislation required every state to adopt. This recent shift in standards-based instruction has persuaded some school leaders to examine and adjust the grading practices that currently exist (O'Connor, 2018). This grading reform effort is focused on aligning grading practices to the standards-based teaching practices that have been adopted (O'Connor, 2018). This change has been met with some resistance from stakeholders (Peters & Buckmiller, 2014).

As educational systems look to grading reform, it becomes essential to include leaders, parents, and students throughout the process. It is imperative that educational leaders have a solid understanding of standards-based grading practices (Townsley, 2019) and realize the deep entrenched beliefs that go along with traditional grading practices. It is recommended that educational leaders move slowly through grading reform; creating a standards-based grading mindset with all stakeholders is a critical component of successful change (Schimmer, 2016). Some research has been conducted to determine the impact of standards-based grading on parents and students, but very little research is available documenting the impact of standards-based grading practices on students. Changing the rules that students have lived by throughout their educational careers could have a negative impact on the academic self-efficacy of students. Now more than ever, students need to persevere as they explore new realms that come with a standards-based grading system.

Chapter III

Design and Methodology

Introduction

The purpose of this study was to analyze the impact of standards-based grading on the academic self-efficacy of high school students. The educational practice of reporting the academic progress of students through grades is a pillar of American education (Vatterott, 2015). Students, parents, and teachers have a right to accurate and efficient feedback regarding a student's areas of academic strength and weakness (Buckmiller et al., 2017). Traditional grading practices fail to accurately communicate student progress and punish students for their mistakes throughout the learning process (Anderson, 2018; Brookhart, 2009; Guskey, 2013; Kohn, 2011). Standards-based grading was developed to address some of the shortfalls of the traditional, norm-referenced grading system (Brookhart, 2009; Marzano, 2000).

This study examined standards-based grading systems through the lens of the impact that this grading system has on the academic self-efficacy of students. Albert Bandura (1977) defined the concept of self-efficacy as a person's belief in their own ability to accomplish a task. This framework has been studied in education and found to be a determining factor in student achievement and student choice regarding educational challenges (Pajares, 1996; Schunk, 1991; Schunk & Pajares, 2002). Students with higher academic self-efficacy choose more challenging activities and demonstrate more academic growth over time (Pajares, 1996; Schunk, 1991; Schunk & Pajares, 2002). Research has shown that self-efficacy perceptions are reduced in students as they progress through school (Pintrich & Schunk, 1996). This decline has been associated with many factors, including traditional grading practices and teacher mindfulness of individual student progress (Schunk & Pajares, 2002).

Self-efficacy has been associated with a student's ability to perform meaningful tasks, persist through challenges, and produce better results when compared to less self-efficacious peers (Schunk & Pajares, 2002). This study looks specifically at the impact of standards-based grading practices on the academic self-efficacy in English class of high school juniors. This chapter provides a description of the specific environment for the study, including the location and selection process for the participants. The study's procedures, including instruments, analysis, and limitations will be addressed in this chapter. Additionally, the design of the research, including methodology procedures utilized to collect and assess data related to individual student academic self-efficacy, will be discussed in detail. This chapter will define the role of the researcher and the specific components involved in this study, including the selection of subjects, location, instrumentation, and analysis. Restrictions to the scope of the study, reliability, and constraints will also be communicated. The following research questions will be addressed from both a qualitative and quantitative research perspective.

Research Questions

The research questions are as follows:

1) What is the impact of standards-based grading on a high school students' academic self-efficacy?
2) What relationships can be drawn between variables associated with the high school student population of schools in a standards-based grading system?

Research Design

The intent of this explanatory, sequential, mixed-methods study was to measure the impact of standards-based grading on the academic self-efficacy of high school students. Due to the complexities involved in correlating student academic self-efficacy to grading practices, both

qualitative and quantitative measures were used. The explanatory sequential design was chosen so that research could be completed in two phases, beginning with the quantitative phase. The second phase used the results from the first phase to develop the qualitative interview protocol (Creswell & Guetterman, 2019). This design allowed the researcher to expose the best of the data from the quantitative and qualitative methods (Creswell & Guetterman, 2019). The instrument and methodology used for the quantitative portion of this study were adapted from the General Self-Efficacy Scale (Appendix A) instrument, which was altered and administered in retrospective post then pre design. This is not to be confused with a pre/post design where subjects take a pre-test before a program begins and then takes a posttest after it is completed. In a post then pre design, subjects are asked about their current experiences or perceptions and then asked to reflect on their prior experiences (Colosi & Dunifon, 2006). This design allowed the investigator to ask the participants to answer questions about their academic self-efficacy before the implementation of standards-based grading and after (Colosi & Dunifon, 2006). The post then pre design allows for the smallest margin of interruption to the classroom setting because the survey is completed all in one setting (Colosi & Dunifon, 2006). Additionally, the foundation supporting this type of research design is that it allows subjects to be tested in one setting with a consistent set of beliefs while eliminating response shift bias (Davis, 2003; Rockwell & Kohn, 1989). Although post then pre design studies eliminate the response shift bias, the researcher must be aware of four specific biases that could become relevant to the validity of a study (Colosi & Dunifon, 2006). These identified biases that a post then pre design is vulnerable to are recall, social desirability, effort justification, and cognitive dissonance (Colosi & Dunifon, 2006). Each of these biases were kept at the forefront during the

construction of the instrument. Furthermore, the study is measuring student perception which makes it appropriate to use the post then pre design model (Colosi & Dunifon, 2006).

Each location for the study was determined through informal interviews with site administration. Principals at four comprehensive high schools, grades 9-12, were asked about the prevalence of standards-based grading practices in their schools. Only the two school sites that indicated the use of standards-based grading practices in English classrooms were used in the study. The first location had all teachers in the English department using a standards-based grading system, while the other school had a mix of teachers, some using traditional grading and others using standards-based grading. The investigator then identified the teachers who were using either standards-based grading ideals or traditional grading practices. Once the classes meeting the standards-based grading criteria were identified, the specific teachers and classes were chosen at random. Teachers were approached regarding participation and asked if they would grant permission to allow their classes to be part of the research. Every subject involved in the study participated voluntarily. Parental consent and minor assent were obtained for all participants.

A sample of students in standards-based grading environments were asked to complete a short survey that asked demographic information and then 10 questions regarding their self-efficacy (Appendix B). The survey was adapted from the Generalized Self-Efficacy Scale, formatting the survey into a retrospective post then pre design. The adapted version of the GSE was piloted with an independent group for both validity and reliability. The students in the study were asked to answer the survey based on their current beliefs about their academic self-efficacy in a class that is using standards-based grading. They were then asked to reflect on their previous, traditional grading experiences and answer the same survey questions. At the end of

the survey, students were asked if they would be willing to participate in a semi-structured interview at a future undisclosed date and time.

The data from all the subjects was compiled, disaggregated, compared, and used to develop clarifying interview questions. The objective of this design was to expound on the quantitative data collected with targeted qualitative interviews (Creswell & Guetterman, 2019). Students indicating a willingness to participate in semi-structured interviews were selected with the goal of having a demographic group that was representative of the overall sample.

This mixed-methods design encompasses the strengths of both qualitative and quantitative data to provide a better understanding of the results to the research questions (Creswell & Guetterman, 2019). The retrospective pre-post design was specifically used to address the vulnerability of the widely used traditional pre then post test design (Colosi & Dunifon, 2006). The semi-structured interview questions helped the researcher clarify results from the survey and gain a deeper understanding of student perceptions regarding the two grading philosophies. The figure below outlines the explanatory sequential mixed methods design.

Figure 8.

Conceptual Process of Explanatory Sequential Mixed-Methods Design

Study participants for this research were identified as 11th grade students currently enrolled in a junior English class with a standards-based grading philosophy. Students from two

large, urban high schools in the Pacific Northwest were used for this study. The schools chosen for this study had recently allowed teachers to voluntarily adopt a standards-based grading philosophy in their classrooms. Student still received letter grades as a final grade, but the standards-based grades they received were based on evidence towards proficiency of the standard rather than an average of performance. Additionally, behaviors such as late work, extra-credit, or attendance were not permissible factors in the student grades.

Initially, the classes selected for the study were chosen based on the grading philosophy in each class. Multiple 11th grade English classes were chosen at each location based on the use of standards-based grading practices. The classes and students chosen for the study were selected in collaboration with site administration and with the willingness of the selected teachers. After the selection process, an email was sent to the teacher to obtain their consent for the study (Appendix I). The students at each school had comparable demographic information, including age, sex, and socioeconomic status (based on the percent of students qualifying for the Free and Reduced Lunch Program or FRLP). The FRLP rates at each school were below 20%. Information regarding consent and demographics can be seen in tables 3 and 4.

Table 3.

Participating School Demographic Breakdown

	School #1	School #2
Schoolwide Free and Reduced Lunch %	10%	17%
Male	54%	52%
Female	46%	48%
11th Grade	22%	23%

The schools in the study were selected for their similar demographics and the availability of standards-based grading platforms. Each school had teachers who had implemented standards-based grading practices within the past three years as well as some teachers who were still using traditional grading practices.

Standards-based grading practices were identified as the following:

- Courses have identified priority standards and proficiency scales in use
- Extra-credit is not given
- Students are given multiple opportunities and options to demonstrate proficiency
- Students' grades are not penalized for behavior issues, including late work
- Final grades are based on evidence of standards mastery rather than an average of the overall scores.

For the purposes of this study, the identifying practices for traditional grading were the following:

- Grades are averaged, norm-referenced grades
- Adopted state standards are not individually reported on through the grading practices

Since the subjects in the study were all minors, both guardian consent and student assent was obtained prior to administering the survey instrument. Fifty-six students from School #1 were offered the opportunity to take part in the study and 37 returned parent consent. At School #2, 45 students were offered the opportunity to take part in the study with 26 students returning the parent consent forms. All the students that returned the parent consent, completed the survey. The total number of surveys collected from both sites was 63.

Table 4.

Participating Schools

	School #1	Offered Consent	Provided Consent	Completed Survey
English 11	Class 1	27	19	19
English 11	Class 2	29	18	18
	School #2	Offered Consent	Provided Consent	Completed Survey
English 11	Class 1	32	16	16
English 11	Class 2	13	10	10
Total	4 Classes	101	63	63

To introduce the study and obtain parent permission, each class was visited and addressed. Each selected class was informed about the study, and it was explained that participation was voluntary with no grade attached to it. Students were told that the study was looking at the impact of the standards-based grading system on their academic self-efficacy. Additionally, students were told that they would be offered the opportunity to take a brief survey asking about their self-efficacy and collecting some demographic information. All students were given guardian consent forms and an incentive to win five-dollar gift cards to local restaurants if completed consent forms were returned. Although incentives were offered as a reward for students that returned their guardian permission forms, the students voluntarily, without reward, participated in the survey.

All students that participated in the survey were also asked if they would be willing to participate in semi-structured focus group interviews. Sixteen students from School #1 and 16 students from School #2 indicated that they would be willing to participate in qualitative interviews. From the students that expressed willingness to participate in interviews, five students were chosen from each school, based on their demographics. One student at School #1

indicated that they were unable to participate due to a previously scheduled conflict, making the total students interviewed at School #1, four. All five selected students from School #2 participated in the interviews. Table 5 displays the gender and pseudonym for each of the students that participated in the interviews.

Table 5.

Participants for Interviews

Pseudonym	Gender	School
Jennifer	Female	Site #1
Ellie	Female	Site #1
Eva	Female	Site #1
Andrew	Male	Site #1
Laura	Female	Site #2
Mary	Female	Site #2
Beth	Female	Site #2
Cory	Male	Site #2
David	Male	Site #2

Data Collection

This study was a mixed-methods, explanatory sequential design that examined the impact of a standards-based grading system on the academic self-efficacy of high school 11th grade English students. Participants in the study were chosen based on their enrollment in standards-based grading classes. Due to the use of human participants, specifically minors as a protected population, protections of confidentiality and guardian consent were obtained prior to using any data in the study (Creswell & Guetterman, 2019; Marshall & Rossman, 2016). Although confidentiality and trust were pillars of this study, the design of the research prevented true

anonymity during the semi-structured interview process. The researcher received certification through the Association of Clinical Research Professionals in the area of Ethics and Human Subject Protection (Appendix D). Permission from the Northwest Nazarene Institutional Review Board was obtained (Appendix E). Finally, permission from the superintendent, district board of trustees, and site principals at each of the schools was also obtained prior to any research taking place (Appendix F, G).

After receiving all approvals, 11th grade English classes at each school were purposefully chosen, and students were given the opportunity to take part in the study. The selection of these classes was purposeful in that, the total sample needed to be composed of only classes that employed standards-based grading philosophies. Site administrators and teachers were contacted to determine the grading practices that were being used in each specific class. Evidence such as grade books, student work, and informal interviews were used to confirm the grading practices.

The quantitative data was obtained through the administration of a survey that was adapted from the GSE (Generalized Self-Efficacy Scale), a tested instrument deemed reliable with Cronbach's alpha between .76 and .90. The GSE is designed to elicit student perceptions on their own abilities to accomplish difficult tasks. Survey questions were not altered from the GSE, but the survey was reformatted into a retrospective post then pre design. Due to the restructuring of the GSE, the new survey was piloted with 26 students that were not chosen to participate in the study. The students were informed of the pilot and given the option to participate. All students opted to participate in the survey. Students in the pilot were asked at the end of the survey for feedback regarding the clarity of the questions and the structure of the instrument. A Cronbach's alpha was run on the pilot study at a .921. Based on confirmation of both validity and reliability of the GSE, the data collection proceeded (Creswell & Guetterman,

2019. The results of the quantitative data obtained were analyzed and drove the semi-structured interviews that were conducted for selected participants.

The qualitative portion of the study was completed through in-person, semi-structured interviews. The quantitative data was used as a roadmap for structuring clarifying questions that would still allow the participant latitude to expand as needed (Marshall, & Rossman, 2016). The qualitative interviews were conducted individually with the five participants being selected from each site at random. Interview questions were structured around the results of the survey for the entire sample size, rather than for the specific participant that was being interviewed. The table below illustrates the number of participants that volunteered to meet in a small group session and the number that were chosen to participate.

Table 6.

Survey Participants and Interview Volunteers

	School #1	Survey Participants	Interview Volunteer	Interviewed
Class 1 Eng. 11	Standards-Based Grading Practices	19	8	2
Class 2 Eng. 11	Standards-Based Grading Practices	18	8	2
	School #2	Survey Participants	Interview Volunteer	Interviewed
Class 1 Eng. 11	Standards-Based Grading Practices	16	10	2
Class 2 Eng. 11	Standards-Based Grading Practices	10	6	3

Quantitative Data

All of the quantitative data collected was collected solely by the author of the study. The researcher visited each class to explain the purpose of the study and offer students the

opportunity to participate. All the students in the classes selected for the study were offered the written consent form for guardians to complete, and students who returned the completed form the following day were given the opportunity to take the survey. As an incentive for students to return the guardian consent, students who returned the completed form were given a raffle ticket and a chance to win a five-dollar gift card to local restaurants. The survey was constructed and offered to all students in the class through Qualtrics, an online survey platform. The researcher received verbal assent from all students prior to administering the survey, and the first question on the instrument asked if they were willing to participate. Any students that indicated they did not wish to participate were automatically forwarded to the end of the survey. Students that did not return the signed permission slips were not given the opportunity to take the survey. The gift cards were offered as an incentive to return the consent forms only, and were not intended to influence survey results. All surveys were administered and completed electronically on devices supplied by the school district. Data from the 63 completed surveys were transferred into SPSS for further analysis.

Qualitative Data

All students that participated in the survey were asked if they would be willing to participate in semi-structured interviews. If they indicated their willingness to participate in these interviews, the instrument asked for the student's information. In total, 32 students agreed to participate in focus group interviews. Of these 32 students, four were selected from School #1 and five were selected from School #2. Students participating in the study at each school were selected based on demographic information and availability. The students were interviewed during lunch and part of the following class. All students that arrived at the interviews were asked for their permission to participate in a recorded interview. These interviews were set up as

focus group interviews in which all students engaged in a dialogue about standards-based grading. The students were asked specific questions but given the opportunity to elaborate if they wanted to or if the researcher needed further information. The interviews were video recorded and transcribed prior to being coded for common themes. Transcripts of the interviews were uploaded into NVIVO to assist with the coding organization and theme recognition.

Analytical Methods

Data analysis was completed for all qualitative and quantitative data. The quantitative data was collected and analyzed to help form the questions for the semi-structured interviews. After the interviews were completed, all data was analyzed together through the theoretical framework of self-efficacy. Through this lens the researcher analyzed the impact that standards-based grading has on a student's belief in his/her ability to complete challenging tasks. Quantitative data were analyzed through scoring of the responses to the Likert scale survey questions and using SPSS to obtain descriptive data. The Likert scale questions were scored using the scoring system:

1- Strongly Agree

2- Agree

3- Disagree

4- Strongly Disagree

The pre-post results were compared against themselves using the Wilcoxon Signed Rank Test in order to determine links and trends within the data. The information identified through the data analysis of the quantitative data was then used to construct the questions for the semi-structured interviews. The qualitative data obtained through the interview process were used to triangulate the information obtained in an effort to answer the research questions stated above.

Limitations

Although this study addresses the impact of standards-based grading on the academic self-efficacy of students, it does have some limitations. The study was purposefully planned to maintain fidelity and consistency though all research. The author recognizes the challenges that are presented in any study and listed these limitations below.

- This study took place at urban schools in the Pacific Northwest area only, in one specific school district. Due to the location and demographic make-up of the schools, the study participants were predominantly Caucasian students from middle-class homes.
- The overall size of the study was also a limitation. Larger sample sizes across varying campuses within the United States would be of benefit in verifying these findings and supplementing this research with additional information.
- The pandemic of COVID-19 created some limitations with grading reform. The researcher found that traditional grading practices had largely been abandoned by the schools in the study. Additionally, participants were hard to contact due to the nature of remote learning. Lastly, it was challenging to obtain guardian consent forms due to the modified schedule at both schools.

Chapter IV

Results

Introduction

In education, grades are assigned to students as a mechanism to communicate evidence of academic progress to all stakeholders (Brookhart, 2009; Guskey, 2015; Guskey & Jung, 2013; Townsley, 2019; Townsley & Buckmiller, 2020). The evidence collected could be formative and summative assessments, classroom assignments, reports, projects, or a combination of these (Brookhart, 2009; Guskey, 2015; Guskey & Jung, 2013; Townsey & Buckmiller 2020). Assigned grades are intended to align with state adopted content standards. Educational content standards are specific statements for each subject area that describe what a student should know and be able to demonstrate at the appropriate level of proficiency (Guskey & Jung, 2013; Schimmer, 2016). Unfortunately, factors such as the use of zeros, extra credit, unbalanced weighting of scores, and penalizing students for behaviors have resulted in inaccurate or misaligned reporting of grades (O'Connor, 2007; Schimmer, 2016; Townsey & Buckmiller 2020). As a result, the grade assigned to an academic outcome does not always accurately communicate a student's academic achievement (Brookhart et al., 2016; Guskey, 2015; O'Connor, 2007). Traditional grading practices that take an average of performance on a variety of work are skewed by non-academic factors and are not advanced enough to accurately report student progress on specific standards (Brookhart et al., 2016; Knight & Cooper, 2019; Townsley, 2019). While educational reform has focused on standards-based instructional practices, grading practices in the United States, dating back to the advent of compulsory education, have been neglected (Guskey & Link, 2019; Starch & Elliot, 1912). As schools across the United States recognize the inaccuracies that traditional grading practices present,

many begin to investigate shifting into a standards-based grading model (Knight & Cooper, 2019; O'Connor, 2018; Townsley, 2019).

Although standards-based grading can look slightly different based on the philosophy of the district, at its core, standards-based grading is defined as a system of reporting student progress specific to the standards that are being assessed (Knight & Cooper, 2019). This system focuses on the most recent evidence of student learning and eliminates outside factors, such as behavior, from the reporting of grades (Iamarino, 2014; Knight & Cooper, 2019; O'Connor, 2018). As schools reflect on grading practices, several have begun to voluntarily move toward a standards-based grading model (Townsley & Buckmiller, 2020).

Educational institutions attempting to facilitate change in the long-held philosophies that frequently accompany traditional grading practices have been presented with challenges by teachers, students, and parents (Townsley, 2019; Townsley & Buckmiller, 2020; Schimmer, 2016). Although some research has been conducted on the transition into a standards-based grading model, the focus has been around three specific areas: parent reactions to the new grading philosophy, the impact on teachers, and implementation from a leadership perspective (Townsley & Buckmiller, 2020).

The purpose of this study was to investigate the impact that a standards-based grading system has on the academic self-efficacy of high school students. Self-efficacy is defined as an individual's perception of their own ability to successfully accomplish a specific task (Bandura, 1997). Self-efficacy is a key component to Albert Bandura's Social Cognitive Theory (Bandura, 1986). The concept of self-efficacy has been linked to educational outcomes, including a student's ability to overcome challenges and attempt difficult tasks in the classroom (Bandura & Schunk, 1981; Hannon, 2014; Honicke & Broadbent, 2015; Komarraju & Nadler, 2013; Pajares,

2009; Pajares & Schunk, 2001). This mixed methods study used the General Self-Efficacy Survey (GSE) (Appendix A) to evaluate the impact of standards-based grading on high school students' self-efficacy. The literature review on self-efficacy provides substantial evidence that a person's level of self-efficacy directly impacts their own work product, the work product of others, verbal interactions, and physiological reactions (Bandura, 1977; Redmond, 2010).

Using Bandura's theory of self-efficacy as the theoretical framework, a post then pre explanatory sequential design was utilized to analyze how grading systems impacted an 11th grade high school student's academic self-efficacy. The research questions were as follows:

1. What is the impact of standards-based grading on a high school students' academic self-efficacy?
2. What relationships can be drawn between variables associated with the high school student population of schools in a standards-based grading system compared to their previous experience in a traditional grading system?

Chapter IV includes an overview of the high schools that were included in this research, the instruments used, validity and reliability procedures and the overall results.

Participant Profile

The participants for this study were chosen using purposeful sampling. This procedure involves the researcher selecting participants based on a specific set of criteria (Creswell & Guetterman, 2019). Participants were identified as part of the larger, targeted sample of approximately 1,350 eleventh grade high school students in the selected geographic location that were currently being assessed using a standards-based grading practice. The following criterion was developed and used to identify potential participants for this study:

- Participants were current 11th grade high school students enrolled in a comprehensive public high school.
- Classes identified for the study were chosen based on subject matter (English Language Arts) and showed evidence of standards-based grading practices.
- Students considered for the study had to be enrolled in one of the identified classes.
- Students involved in the study transitioned to a standards-based grading system within the past three years.

Based on the above parameters, two schools in a large urban school district were selected to participate in the study. Below is a detailed description of the two schools, survey participants, and subjects that were chosen for the semi-structured focus groups.

Vignettes of Schools

The schools chosen for the study were selected first for their geographic locations. Both schools were located in a large urban school district in the Pacific Northwest. After looking at the geographic locations, schools and classes were identified based on the prevalence of standards-based grading practices. The schools and/or teachers in this district had voluntarily moved to a standards-based grading model within the past five years, although these practices are uniform across all schools in the district. Neither of the schools selected to participate in the study had a school-wide policy on standards-based grading practices. The two schools selected had similar demographic information (see Table 7 below).

To conduct the study, the researcher first received approval from the site principal, the district superintendent, and the board of trustees (Appendix F, G). After approval, teachers were contacted through email to obtain voluntary permission for their classes to participate in the study. Upon receiving permission to conduct the study, the researcher confirmed the standards-

based grading practices in the classroom through a short discussion with the teacher and then arranged a convenient time to administer the survey instrument.

Site #1 profile. Site one is classified as a large, comprehensive high school located in an urban school district within the Pacific Northwest region of the United States. The total student population was 2,597, with a reported 84% of students being white, 8% Hispanic, 2% Asian, and 2% two or more races (National Center for Educational Statistics, 2018). All other ethnicities were represented at 1% or less.

Site #2 profile. Similar to Site #1**,** Site #2 is considered a large, comprehensive high school located in the same school district. The total student population was 2,413 with a reported 79% of the student being white, 10% Hispanic, 2% Asian, 2% African American, and 5% two or more races (National Center for Educational Statistics, 2018). All other ethnicities were represented at 1% or less races (National Center for Educational Statistics, 2018). On average, 49% of the student body identified as female and 51% identified as male (National Center for Educational Statistics, 2018). The graduation rate at this location sat above 90%. On average, the students that qualified for the National Free and Reduced Lunch Program is less than 15%.

Participant Profile

Survey participants. The quantitative portion of this study was completed using the General Self-Efficacy (GSE) Instrument as a tool to measure the self-efficacy of individual students. The GSE was administered in a post and then pre design, where students were asked to complete the GSE questionnaire reflecting on standards-based grading practices and then answer the same questionnaire as they reflected on their experiences in a traditionally-graded English class (Colosi & Dunifon, 2006). The students were asked to complete both post and pre test at

the same time. Although parental/guardian consent and verbal student assent was obtained prior to the administration of the GSE, students could stop the survey at any point. Of the 63 students that gave verbal assent and returned guardian consent, all 63 completed the GSE survey.

Table 7 shows the length of time that each student has spent in a standards-based grading system and the combined demographics of all participants in the survey at both sites. It is important to note that standards-based grading was not implemented in either of the schools three years prior to the study. Students who indicated greater than two years in a standards-based grading environment would have no more than three years unless they received formal education outside of their current school. Most of the students that participated in the survey at Site #1 indicated that they had greater than 2 years of experience in a standards-based grading classroom. Exactly half (50%) of the students who participated in the survey at Site #2 indicated that they had greater than two years of experience in a standards-based grading classroom, while the other half indicated they had one to two years of experience with the model. Survey responses and demographic information is indicated below.

Table 7.

Participant Demographics by Location (n=63)

Question and Responses	Site #1	Site #2	Total	Total Percent
Grade Level:				
11th	37	26	63	100%
What is your gender?				
Male	17	14	31	49%
Female	19	11	30	48%
Does not wish to self-identify	1	1	2	3%
What is your race/ethnicity?				
White	32	21	53	84%
Hispanic or Latino	2	2	4	6%
Black or African American	2	1	3	5%
American Indian or Alaska	0	0	0	0%
Native	0	0	0	0%
Asian/Pacific Islander	0	2	2	3%
Other	1	0	1	2%
Does not wish to self-identify	0	0	0	0%
Parents education level:				
Mother High School Diploma:				
Yes	36	25	61	97%
No	1	0	1	1.5%
Not Sure	0	1	1	1.5%
Father High School Diploma:				
Yes	36	22	58	92%
No	1	3	4	6%
Not Sure	0	1	1	2%
Mother College Degree:				
Yes	24	15	39	62%
No	11	9	20	32%
Not Sure	2	2	4	6%

Father College Degree:				
Yes				
No	10	8	18	29%
Not Sure	26	14	40	63%
	1	4	5	8%
Length of time student has been in a Standards-Based Grading English class:				
0-1 Years	3	0	3	5%
1-2 Years	8	13	21	33%
Greater than two years	26	13	39	62%
Percent of total sample size by length of time in a Standards-Based grading English class:				
0-1 Years	8%	0%		
1-2 Years	22%	50%		
Greater than two years	70%	50%		

Interview Participants

At the end of the quantitative survey, participants were asked if they would be interested in meeting with the researcher on a future, undisclosed date to discuss their experiences with standards-based grading. At Site #1, sixteen out of thirty-seven agreed to participate in a semi-structured focus group, and at Site #2, sixteen out of twenty-six agreed to participate. Of the students who indicated their willingness to take part in the interviews, five students at each site were randomly selected and contacted to participate. The selected students met in a large conference room at each site and discussed various topics around standards-based grading. Involvement in the semi-structures focused groups was voluntary. The interviews were video recorded and transcribed for qualitative coding and analysis.

Table 8

Focus Group Participant Demographics

Pseudonym	Gender	School
Jennifer	Female	Site #1
Ellie	Female	Site #1
Eva	Female	Site #1
Andrew	Male	Site #1
Laura	Female	Site #2
Mary	Female	Site #2
Beth	Female	Site #2
Cory	Male	Site #2
David	Male	Site #2

Survey Validity and Reliability

Validity and reliability were at the forefront in the design of this study. Validity refers to evidence that the analysis and conclusions of the survey were in alignment with the purpose of the survey (Creswell & Guetterman, 2019). The design of this study was a post then pre retrospective design, in contrast to the pre/post design. This design allowed the subjects to respond to a survey regarding their current perceptions and then reflect on their perceptions prior to standard-based grading, compared to the pre/post design which is completed in separate settings. The post then pre design, or the retrospective pretest, eliminates the validity critiques that frequently come with a pre/post design where it is difficult to avoid response shift bias (Colosi & Dunifon, 2006). Since subjects take the survey all in one setting as they are being asked about their own perceptions, the post then pre design allows for the most valid responses (Colosi & Dunifon, 2006). Reliability refers to the ability of the results of the survey to be

replicated in similar environments (Creswell & Guetterman, 2019). Prior to any data collection, the researcher took specific measures to ensure the reliability and validity of the study.

Survey pilot. The General Self-Efficacy (GSE) instrument is a published scale with an internal reliability with Cronbach's alphas between .76 and .90 (Schwarzer & Jerusalem, 1995). Since the survey was altered slightly and formatted into a post then pre retrospective design (Appendix B), a pilot was run with a group of non-participating students of similar demographics. All of the students who participated in the pilot were 11th grade students enrolled in an 11th grade English class using a standards-based grading philosophy.

Cronbach's Alpha of pilot survey. The results of the pilot survey were analyzed to verify internal consistency by conducting Cronbach's alpha. This test was used to measure the results of the questionnaire regarding the consistency of the responses (Creswell, 2015; Field, 2018; Maxwell, 2012). With the published reliability of the GSE scale at .76-.90 (Schwarzer & Jerusalem, 1995), the researcher used these parameters to analyze the reliability of the modified survey. Since the survey was conducted in the post-pre retrospective design, the results were analyzed independently (Colosi & Dunifon, 2006). The Cronbach's alpha results for the modified GSE were specific to students before standards-based grading was analyzed in its entirety. Then, Cronbach's alpha was determined for the GSE after standards-based grading. Each set of survey results was also analyzed, per question. A total of 36 students completed the instrument tool. The pilot study produced an overall internal reliability Cronbach's alpha level of .921 on the pre-survey and a Cronbach's alpha level of .849 on the GSE after the implementation of standards-based grading. Looking specifically at each individual question, on the pre-survey data, one question was identified for possible removal in order to increase the instrument's reliability. Removing question #2 two of the survey would have increased the

reliability to .933 from a .921. Because the reliability of the pre-survey met all reliability parameters, all questions were kept. On the post survey, the Cronbach's alpha was slightly lower, .849 as compared to .921, but still fell well within the reliability recommendations. The question-by-question analysis did not show any extreme outliers that would impact the reliability. Therefore, all questions were left in the post survey. Based on the Cronbach's alpha of both instruments, the post-pre retrospective survey was determined to be reliable at .921 prior to SBG and .849 after SBG. Table 9 illustrates the final reliability data for both sets of survey results.

Table 9.

Cronbach's Alpha Reliability

	N	Cronbach's Alpha	Item Deletion High	Item Deletion Low
GSE Reliability prior to SBG	36	0.921	0.933	0.904
GSE Reliability after to SBG	36	0.849	0.85	0.823

Additionally, after the survey was completed students were asked if there was anything confusing about the survey, any verbiage that was not understood, or any challenges that the researcher needed to address prior to future administrations. Based on student responses and the Cronbach's alpha reliability, there were no changes made to the modified survey instrument.

Semi-Structured Focus Group Pilot Validity and Reliability

The qualitative portion of this study, which was conducted through focus group interviews, was also piloted. The purpose of this pilot was to ensure that the interview questions elicited the information needed to support the research questions proposed. The interview pilot was conducted with one focus group of five students that did not take part in any portion of the

study. These participants were selected on a voluntary basis from the students who completed the pilot survey. The pilot interviews were recorded and reviewed by the researcher. Based on the results of this analysis, there were no changes made to the interview questions. Table 10 provides the demographic information of the participants in the pilot interview.

Table 10.

Demographics of Pilot Interview Participants

	Gender	Year in School
Pilot Student 1	Male	11th Grade
Pilot Student 2	Female	11th Grade
Pilot Student 3	Female	11th Grade
Pilot Student 4	Female	11th Grade
Pilot Student 5	Male	11th Grade

Data Collection Instruments

Quantitative survey instrument. The survey instrument used for this study was divided into three sections. The first section asked students for their assent to take the survey and then asked for the length of time the student had been involved in a standards-based grading model and their current grade level. Since the survey was only administered in 11th grade ELA classes, all participants were 11th graders.

In the second section, students were asked questions about their self-efficacy using Schwarzer and Jerusalem's (1995) General Self-Efficacy Scale. The GSE survey was slightly modified by the researcher into a post then pre design (Appendix B). This portion of the survey asked students to answer ten questions about their self-efficacy in an English class that was graded in a standards-based model. The students were then asked to reflect on previous experiences in English classes that were graded in a traditional model and answer the same ten

question survey. The questions for both surveys were on a four-point Likert scale. The scale had the following rankings:

- 1 = Not at all
- 2 = Hardly True
- 3= Moderately True
- 4 = Exactly True

The final score for each survey was calculated by finding the sum of all the items. For each survey, scores ranged from ten to forty. Results from the two surveys were calculated independently and then compared to determine growth or diminished levels of self-efficacy.

The third section of the survey instrument asked students about their demographic information and about the education level of their parents. Specifically, students were asked about their gender and ethnicity, then they were asked about the education level of their mother and father. These demographic questions were asked to assist the researcher during the analysis phase of the study to determine if there were any correlations between demographic information and the results on the GSE scale. The results of the quantitative portion of this study will be discussed later in this chapter.

Qualitative interview protocol. The interview questions for the qualitative portion of this study were created by the researcher to guide the semi-structured focus groups. All participants were read the same script prior to the interview taking place (Appendix H). At the culmination of the interview, each student was sent a copy of their transcribed responses, via email, to ensure the students were accurately represented.

Quantitative Results

The explanatory sequential design recommends the researcher analyze the quantitative data prior to and independently of the qualitative data analysis (Creswell & Guetterman, 2019). This process allows the researcher to use information obtained through the quantitative data analysis to guide the semi-structured interviews. In an effort to analyze the quantitative data, the researcher used IBM SPSS 27 to look at the assumptions needed to complete the Shapiro-Wilk test. The assumption of normality in the data set was not met, and, therefore, non-parametric measures were utilized (Field, 2018). The Wilcoxon signed-rank test was used to analyze Research question #1 in order to determine the overall impact of standards-based grading on the academic self-efficacy of students. The Wilcoxon signed rank test is the non-parametric version of the paired sample t-test and is used by researchers to compare two set of data that come from the same set of participants, before and after an event has occurred (Field, 2018). The Wilcoxon signed-rank test was also used to analyze research question #2. Each individual question on the GSE scale was evaluated to determine the impact of standards-based grading. The results of this analysis were reported through the mean ranking, *p*-value (statistical significance), *z*-score, and effect size (Field, 2018). *Z*-scores are determined from the Wilcoxon signed-rank test and display how many standard deviations a score is from the mean. Statistical significance, as displayed by a *p*-value, tells the researcher the reliability of the finding (Field, 2018). Effect size is used to determine the size or magnitude of the difference between two groups (Field, 2018). The effect size was calculated using Cohen's d. The researcher used the following breakdown to report on the effect size.

- 0.2 small effect size
- 0.5 medium effect size

- 0.8 large effect size (Field, 2018)

The researcher also analyzed the quantitative data for each research question, both how the data related to gender and the amount of time each student had spent in a standards-based grading classroom.

Quantitative Results for Research Question 1: Standards-Based Grading Influence on Academic Self-Efficacy

The influence that academic self-efficacy has on a student's choice in difficult tasks, perseverance through challenge obstacles in learning, and overall academic success has been well documented (Bandura & Schunk, 1981; Hannon, 2014; Honicke & Broadbent, 2015; Komarraju & Nadler, 2013; Pajares, 2009; Pajares & Schunk, 2001). Additionally, research has shown that educational institutions continue to receive pushback from teachers and parents as they work to transition into this new grading philosophy (Townsley & Buckmiller, 2020). The intent of research question #1 was to look at the overall impact standards-based grading had on the academic self-efficacy of 11th grade students. The question was written as follows: "What is the impact of standards-based grading on a high school student's academic self-efficacy?" To answer this question, the researcher used a post-pre retrospective survey based off the GSE scale.

A Wilcoxon signed rank test was conducted to determine the effect of standards-based grading on academic self-efficacy. Sixty-three participants were selected to complete the GSE based on their perceptions both prior to standards-based grading and after the implementation of standards-based grading. Of the sixty-three participants, thirty-six (57%) reported an increase in academic self-efficacy, twelve (19%) stayed the same, and fifteen (24%) had decreased academic self-efficacy. The overall GSE score from students based on their experiences prior to standards-based grading had a mean of 31 with a maximum of 40, a

minimum of 17, and a standard deviation of 5.39. As compared to the overall GSE score from students based on their current experiences in a standards-based grading environment which yielded a mean score of 32.78, with a maximum of 40, a minimum of 10, and a standard deviation of 5.24.

There was a statistically significant difference ($p < .05$) in overall self-efficacy of 11th grade students after the implementation of standards-based grading (Median=33.00, Mean = 32.78) as compared to traditional grading (Median 31.00, Mean = 31.0), $z = -2.515$, $p = .012$. Effect size was calculated using Cohen's d at d= .249, which indicated a small effect size between the two mean scores (Field, 2018). An effect size is a measure of magnitude, and table 11 provides a detailed look at the analysis for overall GSE scores of high school students' self-efficacy.

Table 11.

Impact of Standards-Based Grading on Self-Efficacy of Students

	N	Mean before SBG	Mean after SBG	Median Before SBG	Median After SBG	p Value	Z-Score	d Effect Size
What was the overall General Self-Efficacy score?	63	31	32.78	31	33	0.012	-2.515	0.224

The overall GSE scores were also analyzed through the demographic information that was specifically collected on the survey instrument. Due to the sample size being unevenly distributed toward students that identified themselves as Caucasian (84%), ethnicity indicators

were reviewed but not reported. Additionally, the parent/guardian education level was reviewed but not reported on because most of the participants indicated that both of their parents received at least a high school diploma.

The GSE scores results were calculated for gender with n = 30 (%) students identifying as female, n = 31(%) students identifying as male, and with two students choosing not to identify. The two main subgroups of male and female were analyzed in regard to overall self-efficacy scores on the GSE. There was no statistically significant difference (p=.089) found on the impact of self-efficacy in a standards-based grading system for females (Median=33.00, Mean = 30.8667) and males (Median=33.00, Mean = 32.7097). There was a statistically significant difference (p=.046) found on the impact of self-efficacy in a standards-based grading environment as compared to traditional grading experiences in males (Median=33.00, Mean = 32.7097) and females (Median 31.00, Mean = 31.1935). Lastly, GSE score were disaggregated by amount of time that students indicated being in a standards-based grading classroom, which yielded no significant difference (p=.1, .078, and .285). Although increases in mean ranking were observed in all subgroups when compared to prior experiences in traditional grading platforms, the subgroups provided very little significant findings. Table 12 shows the overall scores as they were analyzed through demographic information.

Table 12.

Survey Results for Research Question 1 by Gender

Female	**N**	**Mean Before SBG**	**Mean After SBG**	**Median Before SBG**	**Median After SBG**	**p Value**	**Z-Score**	**d Effect Size**
What was the overall General Self-Efficacy score?	30	30.8667	33.133	31.5	33	0.089	-1.703	-0.3109
Male	**N**	**Mean before SBG**	**Mean after SBG**	**Median Before SBG**	**Median After SBG**	**p Value**	**Z-Score**	**d Effect Size**
What was the overall General Self-Efficacy score?	31	31.193	32.7097	31	33	0.046	-1.996	-0.3585
Survey Results for Research Question 1 by length of time in SBG								
0-1 Years	**N**	**Mean before SBG**	**Mean after SBG**	**Median Before SBG**	**Median After SBG**	**p Value**	**Z-Score**	**d Effect Size**
What was the overall General Self-Efficacy score?	3	28.6667	34.33	30	34	0.285	-1.414	-0.8164
1-2 Years	**N**	**Mean before SBG**	**Mean after SBG**	**Median Before SBG**	**Median After SBG**	**p Value**	**Z-Score**	**d Effect Size**

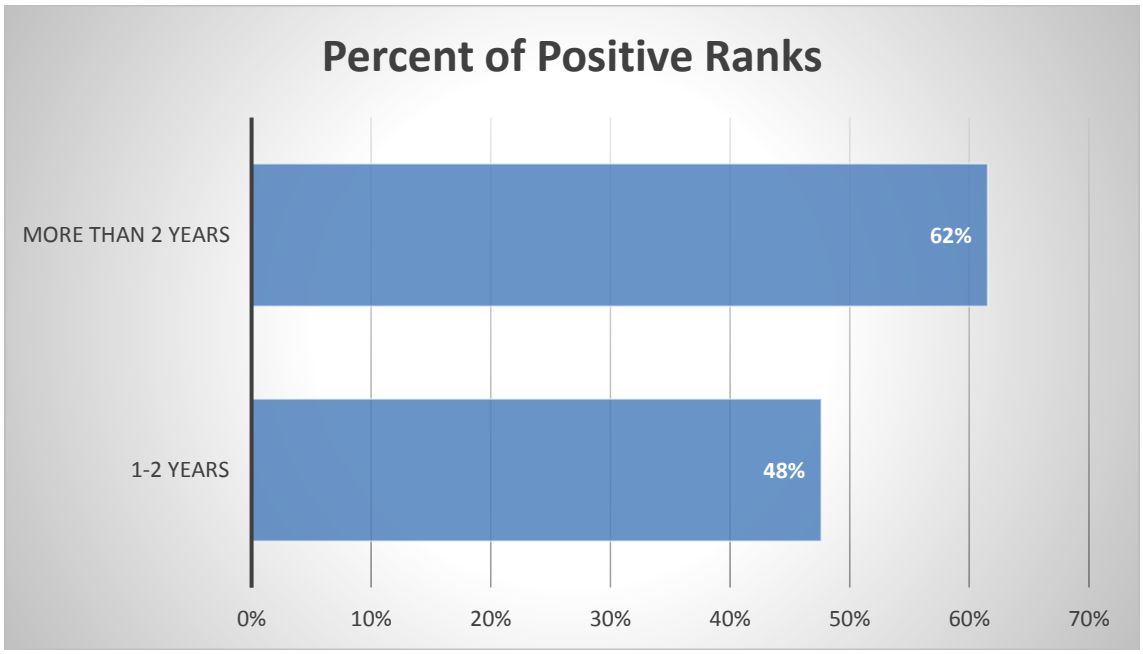
What was the overall General Self-Efficacy score?	21	29.8095	31.5714	30	30	0.078	-0.447	-0.0975

Greater than 2 Years	**N**	**Mean before SBG**	**Mean after SBG**	**Median Before SBG**	**Median After SBG**	**p Value**	**Z-Score**	**d Effect Size**
What was the overall General Self-Efficacy score?	39	31.8205	33.3077	32	34	0.1	-0.119	-0.0191

In addition to the measures of central tendencies, the General Self-Efficacy scale scores were also analyzed in terms of positive and negative movements on overall score. Students indicated a 14% increase in amount of positive ranks given overall on the GSE from 1-2 years as compared to more than 2 years in an SBG setting. Due to the small number of students (n=4) that indicated working in an SBG classroom for 0-1 years, the results were left out of this analysis. Figure 9 shows the percentage of positivity ranks going up as students spent more time with the standards-based grading philosophy.

Figure 9.

Percentage of Positive Ranks



Qualitative Results for Research Question 2: Relationship between Each Specific Question on the GSE.

The GSE scale was comprised of ten Likert scale questions that are valued and combined for one overall score. The intent of research question #2 was to look at each individual question to determine the effect a standards-based grading philosophy had as compared to a traditional grading platform. Understanding how each element of the GSE is impacted could provide valuable insight into the impact of grading practices on students. The Wilcoxon signed-rank test was used to analyze the data for this research question. According to student responses, every question asked had an overall mean increase as students went from traditional grading to

standards-based grading. There was a statistically significant difference found between five of the ten questions on the GSE instrument. These questions are:

- I am confident that I could deal efficiently with unexpected events. (Z=-3.118, p=.002, d= .28). Before SBG (m= 2.79). After SBG (m=3.27).
- Thanks to my resourcefulness, I know how to handle unforeseen situations. (Z=-2.119, p=.034, d= .19). Before SBG (m= 3.03). After SBG (m=3.24).
- I can remain calm when facing difficulties because I can rely on my coping abilities. (Z=-2.142, p=.032, d= .19). Before SBG (m= 2.92). After SBG (m=3.16).
- When I am confronted with a problem, I can usually find several solutions. (Z=-2.353, p=.019, d= .21). Before SBG (m= 3). After SBG (m=3.29).
- If I am in trouble, I can usually think of a solution. (Z=-2.267, p=.023, d= .20). Before SBG (m= 3.16). After SBG (m=3.38).

The ten questions are listed in the table 13, below, along with their p-value, z-score, median value, mean ranking, and effect size.

Table 13.

Survey Results by Question

		N	Mean before SBG	Mean after SBG	Median Before SBG	Median After SBG	p Value	Z-Score	d Effect Size
1	I can always manage to solve difficult problems if I try hard enough.	63	3.17	3.22	3	3	0.531	-0.626	-0.0558
2	If someone opposes me, I can find the means and ways to get what I want.	63	2.97	3.05	3	3	0.452	-0.751	-0.0669
3	It is easy for me to stick to my aims and accomplish my goals.	63	3.11	3.21	3	3	0.506	-0.664	-0.0592
4	I am confident that I could deal efficiently with unexpected events.	63	2.79	3.27	3	3	0.002	-3.118	-0.2778

5	Thanks to my resourcefulness, I know how to handle unforeseen situations	63	3.03	3.24	3	3	0.034	-2.119	-0.1888
6	I can solve most problems if I invest the necessary effort.	63	3.48	3.56	4	4	0.382	-0.875	-0.078
7	I can remain calm when facing difficulties because I can rely on my coping abilities.	63	2.92	3.16	3	3	0.032	-2.142	-0.1908
8	When I am confronted with a problem, I can usually find several solutions.	63	3	3.29	3	3	0.019	-2.353	-0.2096
9	If I am in trouble, I can usually think of a solution.	63	3.16	3.38	3	3	0.023	-2.267	-0.202
10	I can usually handle whatever comes my way.	63	3.37	3.41	3	3	0.597	-0.528	-0.047

Qualitative Research Findings

The explanatory sequential design that was used for this study allowed the researcher to use the qualitative data to help clarify any of the quantitative findings. The semi-structured focus group questions were developed prior to quantitative analysis, but then they were reformed to address the specific information needed to clarify the quantitative findings. The semi-structured interview process allowed for the researcher to ask a set of pre-determined questions and then clarify with follow-up questions. All of the interviews were completed in small focus groups, and each session was recorded for audio and video documentation (Creswell & Guetterman, 2019). Notes were taken throughout the interview so the researcher could reference them as needed (Creswell & Guetterman, 2019). The audio portion of the recordings was transcribed by the researcher, and the files were uploaded into NVivo for initial coding purposes. To begin the coding process, the researcher read through the transcripts of both small group interviews (Creswell & Guetterman, 2019). During this process, the transcripts were highlighted for various codes that were observed (Saldana, 2016). The researcher identified 16 codes within the data. NVivo was used to identify key words that were used consistently throughout the interviews. The commonly used words were placed in a word cloud, shown in Figure 10, through NVivo to help begin the coding process.

Figure 10.

Word Frequency

The researcher then read through the documents a second time to place the identified codes into clusters and, finally, into themes (Marshall & Rossman, 2016). Table 14 illustrates the four themes that were found and the key words that were used to identify each theme.

Table 14.

Key Words for Qualitative Themes

Learning Environment	Grading Perceptions	Self-Efficacy	Parents
• Repeatability • Make-up • Possibility • Chance • Redo • Fix • Opportunity • Hard to fail • Challenging • Retake • Prove yourself • Correct mistakes • Increased learning • Less stress • Effort	• Fair • Accurate • Finally master it • Your best • Struggle • Constantly learning • Positive • Clarity	• Trying hard • Safety net • Risks • Get better • Try my hardest • Find a way to do better • Easy transition	• They understood • Communication • Explain to them

The above themes were derived through the coding process of the qualitative interviews. Some of the themes were more pronounced than others throughout the interviews. Table 15 provides a breakdown of each theme and the frequency in which the theme was presented throughout all of the interviews, combined. Each of these themes was discovered through reviewing the transcripts where students were specifically asked about their experiences in a standards-based grading environment as compared to previous experiences in a traditional grading classroom.

Table 15.

Frequencies of Themes

Theme	Frequency
Learning Environment	42
Grading Perceptions	40
Self-Efficacy	24
Parents	16

Discussion of Emerging Themes and Findings

Learning environment. During the small group student interviews, one of the topics that continued to surface was the concept of the learning environment. Students spoke of how the standards-based grading philosophy gave them the safety net that was needed when taking on challenging tasks. The concept of the learning environment surfaced 42 times throughout the small group interview sessions, although none of the questions specifically asked about learning environment. One example is seen in Corey's statement, "…if you fail an assignment early in a traditional grading system, then you're probably in trouble for the rest of the quarter. Whereas in standards-based grading, you can always make that up." Laura followed up with, "It's like, when you get like a C or below, well, for me at least, it's like I always want to go and fix that to get like an A or a B." This idea that learning is a process of mistakes that can be fixed and resubmitted for further evaluation without penalty was viewed a positive aspect of standards-based grading.

Students also discussed how the standards-based grading practices encouraged them to take some additional risks when they were outlining the tasks they plan to work on. When asked specifically about the educational risks that are taken, Ellie explained:

> Yeah, um now I definitely take a lot more risks with what I'm writing just because I know like, if it doesn't work out, then I can just revise it and make it how it should be. Um, but like, especially with English, I kind of always do my own like, little tweak on it, um, because English has always been my like thing, um but, like I was saying earlier, you don't know what the teacher's kind of like or dislike…

Eva followed up this statement, saying:

> I agree with that, I think like, well, English has never really been my thing I've always sucked at writing and so this way I can try and find a way to get better at what I'm doing without having to worry about ok, if I do this and it doesn't work out, it just, that's it, like it's going to affect my grade a lot. Whereas like now, do that and be like, ok, if it doesn't work out, I have a chance to go back and fix it and like, leave it to the way it was supposed to be or talk to the teacher about the way it should be and then fix it from there.

The last part of this learning environment that surfaced throughout the interviews was the idea that the standards-based grading philosophy reduces stress on each student by limiting the impact that an individual assignment has on a grade. It also encourages students to fix small mistakes and continue their learning regardless of the grade that they receive on an assignment. Eva explained:

> I think it actually relieved stress...having the opportunity to re-turn everything in, because like you were saying (pointing to Andrew), it's the dumb mistakes that you don't realize in the moment when you turn it in, that will bring your grade down the most so being able to be like, "Oh that was really stupid, can I fix that?" ... and they'll be like, "Yeah," and then I get full points and improved grades and I feel like it leads to like, so much less stress.

The theme of learning environment presented itself throughout the interviews. Students spoke about the reduced stress, increased learning opportunities, and their willingness to take on challenging educational work. Much research has been done on the benefits of having a positive learning environment for students, and these semi-structured interviews captured student voices expressing the impact of SBG on the overall learning environment.

Grading perceptions. Another theme that surfaced almost as much as learning environment was the concept of grading perceptions. Throughout the interviews, students articulated that they felt the standards-based grading philosophy was fair and provided a more accurate representation of their learning. Students provided comments such as:

- "Now I feel like it's calculated on your like, last three grades, so like basically you could, fail every assignment, but if you do good on your last ones, that's what counts."
- "Because the assignments are focused on the standards, I know exactly what I need to do to demonstrate that I learned something. It is very clear."
- "Or say I struggled at the beginning and then at the end I finally brought my grade up to an A, then that's where it pays off."
- "You're constantly learning the same thing and can master it."
- "I like it a lot because my grades are better now and um, it's been really nice because I'll do it wrong, or sometimes I'll get some points off but then I'll talk to the teacher and it'll be like, "Oh you missed this and this," and I'll just add it in so that makes it good."
- "I feel like the grades are real, I mean if I get a "B" I probably deserve it or if I get an "A" I probably earned it. My grades really show what I have learned."

Although the themes of learning environment and grading perceptions can be very closely related, the researcher differentiated between the two by looking at the actual grading practice

versus the impact of the practice in a learning environment. Standards-based grading practices gave students a feeling of fairness and provided clarity around what needed to be completed to demonstrate competency on a specific standard.

Self-efficacy. The theme of self-efficacy was identified as the interviews were analyzed in conjunction with the results of the GSE survey. Many of the topics that were addressed on the survey began to surface during the semi-structured interviews. This included students dealing with challenging situations, adapting to change, and solving difficult problems. Students specifically addressed self-efficacy related topics 24 times throughout the interviews. Student responses related to self-efficacy include:

- "I choose difficult things to do because I want to learn more."
- "This was crazy to start but now that I understand it, I think it is better, I really believe in my ability to deal with change now."
- "Like, now I feel like I'm a good student and I tried my hardest to keep up and I was able to get good grades and I feel like if it was traditional, or the other form of grading where you were, there's no redos, no late work, that kind of stuff, I would have been falling behind and it wouldn't necessarily have been my fault."
- "Especially in the beginning. Like where are all these assignments I'm doing but not being counted? But then I figured it out and it really pushes me."
- "I feel like transitioning was pretty easy. It, it was a relief because, um, just because I was a... I was a bad student, kind of. I was, it was just very hard to stay motivated and so I would often turn things in late, or turn things in roughly and the...I don't know if this is necessarily a positive or a negative with the new grading system, but because of that, since I was such a like poor student, I was able to still keep up high grades, but then

again, I still learned and I still, like, I wasn't doing things on time. I wasn't doing things properly, but I still got good grades because I was able to go back and redo it, it forces me to work hard rather than just give up."

Self-efficacy has been connected to student motivation and student choice, specifically when it comes to difficult tasks. The focus groups provided evidence that the standards-based grading philosophy helps students deal with change, encourages them to take on difficult tasks, and provides some motivation for them to complete assigned work. The sentiment gained through all of these interviews was that the standards-based grading system provided students with some optimism because they felt there was always a chance to recover from a poor grade in a class.

Parents. Although parents were only specifically discussed at one point in the interview, the influence that they had on the students was evident. Students brought up the challenges presented to them in communicating overall progress and new grading methods. Students explained that this communication gap was one of the most challenging aspects of SBG, with one student stating, "if we didn't understand it, how were we supposed to explain it to our parents?" Although this theme had a smaller representation, with 16 codes, it was a component that came up throughout the interviews. Students explained:

- "My dad still is like, "I don't know what's going on.""
- "I just, like, remember doing one assignment, it was, I think it was a paper, and I did it, and I turned it in. And then, like, we did another one and my mom was like, "Why didn't this one count towards your grade?" and I honestly didn't know."
- "I remember this being a conversation at dinner for the entire first year, my parents would say this is stupid, I would try to explain and they would still say, this is stupid."

- "My parents were really mad, because I had an F then I had an A. It is weird like that."

Conclusion

Chapter IV provided a detailed description of the demographic information for the participants, data collection methods, validity and reliability verifications, and a review of the quantitative and qualitative findings. The design of this study was explanatory sequential mixed methods, including both quantitative and qualitative data analysis. The quantitative data was collected using the post then pre retrospective design with the GSE survey as the instrument. This design allowed the participants to answer questions based on their current beliefs and then answer the same questions based on their beliefs in a prior grading system. All of the survey questions were scored and calculated on a Likert scale.

The first research question addressed the impact of standards-based grading on the overall academic self-efficacy of students. This research question was addressed through the analysis of total scores on the GSE survey. The total scores showed a statistically significant increase in GSE scores for students after the implementation of standards-based grading. The qualitative data expounded on and gave further support to this finding. The second research question looked at the individual indicators within the GSE survey. The findings substantiated an increase in the mean of all scores with a statistically significant increase in five of the ten questions on the GSE. The overall results are discussed in detail in Chapter V.

Chapter V

Discussion

Introduction

Grading has become a foundational pillar of American educational institutions, where there is strong support for traditional practices and great apprehension toward change (Vatterott, 2015). The first factor that educators must determine as they analyze their grading practices is the purpose behind their grades (Guskey & Jung, 2013). Research shows that educators put the purpose of grading into the following six categories: 1) communicate academic progress to parents, 2) communicate academic progress to students, 3) identify students for specific educational programs, 4) incentivize students, 5) evaluate educational programs, and 6) provide feedback regarding student effort (Airasian, 1997 Brookhart, 2009; Frisbie & Waltman, 1992; Guskey, 2015; Guskey & Bailey, 2001). Simply put:

> The purpose of grading is to describe how well students have achieved specific learning expectations based on evidence gathered from an assignment, assessment, or other demonstration of learning. Grades are intendent to inform parents, students, and others about learning success and to guide improvements when needed. (Guskey & Jung, 2013, p. 71)

Unfortunately, grades have become so complex and misguided that they have almost become a meaningless source of communication to stakeholders (O'Connor, 2007; Marzano, 2000; Schimmer, 2016; Townsey & Buckmiller 2020).

The implementation of standards and published expectations of learning outcomes in k12 education has established clear guidelines for what students should know and be able to do in every grade and in every subject (Guskey & Jung, 2013; O'Connor, 2007; Townsley, 2019).

Research has shown, however, that using traditional grading practices to evaluate and effectively communicate student progress towards their proficiency of these standards is ineffective (Guskey & Jung, 2012; Kunnath, 2017, O'Connor, 2018). Despite research supporting grading reform, grading practices have remained largely untouched as schools continue to use outdated systems to report student achievement (Guskey, 2015; Townsley & Buckmiller 2020).

In response to the growing evidence regarding deficiencies present in traditional grading systems, the standards-based grading philosophy was developed (Iamarino, 2014; O'Connor, 2018; Townsley, 2019). Standards-based grading is a system of grading that reports students' progress according to their proficiency on each specific standard (Guskey & Jung, 2013; Townsley, 2019). This grading platform eliminates external factors, such as behavior and extra-credit, with the intent of reporting more accurate grades to all stakeholders (Guskey & Jung, 2013; O'Connor, 2007; Schimmer, 2016; Townsley, 2019). In this model, students are given multiple opportunities to show evidence of proficiency on a standard without penalty (Guskey & Jung 2013; O'Connor, 2018; Schimmer, 2016; Townsley, 2019). When students are afforded multiple opportunities to demonstrate proficiency on a specific standard, they are held to a higher level of accountability (Wormelli, 2011).

Schools that are looking to transition to a standards-based grading system have been met with some resistance from parents due to their discomfort in this new grading model (Peters et al., 2017; Townsley, 2019). Despite the endorsement that standards-based grading has received from veteran teachers (Tierney, Simon, & Charland, 2011), parents have expressed many concerns regarding this new philosophy (Townsley, 2019). The greatest concern expressed by parents involves the impact of this new grading system on post-secondary opportunities for students (Peters & Buckmiller, 2014). With a limited number of peer-reviewed studies

correlating the impact of standards-based grading on standardized test scores (Pollio & Hochbein, 2015; Welsh, D'Agostino, & Kaniskan, 2013), the impact of this grading system is not yet clearly defined (Townsley, 2019). Research does support that standards-based grading improves communication of academic progress through detailed reporting of student proficiency on each academic standard (Brookhart, 2009; Guskey, 2001; Heflebower & Marzano, 2011; Schimmer, 2016).

Academic self-efficacy is defined as a person's belief in their own ability to successfully complete an academic task (Bandura, 1986, 1993, 1997; Dweck, 1986; Honicke & Broadbent, 2015; Pajares, 1995; Schunk & Pajares, 2002). A person's level of academic self-efficacy has been correlated with academic achievement, choice in challenging activities, motivation, and perseverance (Bandura, 1986, 1993; Dweck, 1986; Honicke & Broadbent, 2015; Pajares, 1995; Schunk & Pajares, 2002). If a person has the appropriate skills to perform a task, their self-efficacy has been shown to be a significant factor in predicting success (Bandura, 1977). With limited research correlating standards-based grading to standardized test scores, this study analyzed this grading philosophy through measuring the effect that it has on the academic self-efficacy of 11th grade high school students.

The purpose of this study was to examine the impact of a standards-based grading system on the academic self-efficacy of high school students. The questions examined through this mixed method pre then post retrospective design study included:

1) What is the impact of standards-referenced grading on a high school student's academic self-efficacy?
2) What relationships can be drawn between variables associated with the high school student population of schools in a standards-based grading system?

Each of these questions was measured quantitatively through the GSE scale and qualitatively through small group, semi-structured interviews. Chapter V will provide an analysis and interpretation of the overall results specific to each research question, including a connection between this research, current literature, and the theoretical framework. In addition, recommendations from the researcher for extended studies in grading reform and the implications of this work on current practices will be examined.

Summary of Results

The purpose of this mixed methods study was to analyze the impact of standards-based grading on the academic self-efficacy of high school students. The GSE scale was used to measure student self-efficacy levels both before the implementation of standards-based grading and after. This was done using a post then pre retrospective format, which allowed students in standard-based grading classes to reflect on their current self-efficacy levels and then to reflect on self-efficacy perceptions prior to the implementation of standards-based grading. This method allowed the researcher to collect the quantitative data all in one sitting, eliminating time restraints and limiting response shift bias (Colosi & Dunifon, 2006). An explanatory sequential design was used to collect the data in two separate phases. The first phase consisted of the quantitative data collection and analysis, and the second phase involved the qualitative interviews which were designed to provide further clarity and understanding of the quantitative data (Creswell, 2015). In the quantitative phase of data collection, 63 high school juniors were asked to complete a survey regarding their current level of self-efficacy and then to answer the same instrument regarding their self-efficacy in a traditional grading system. The survey focused on the following components:

- overcoming challenging problems
- persistence towards long-term goals
- solutions focus
- managing difficult situations

Quantitative analysis of the survey data was completed through SPSS, Version 27. A Wilcoxon Signed Rank, non-parametric test was used for analysis (Field, 2018). This test analyzed the positive and negative differences in the GSE scores both before and after the implementation of standards-based grading (Field, 2018).

Following the survey, students were asked if they were willing to participate in a semi-structured, small focus group interview. Of the 63 students who completed the survey, 31 students agreed to participate. This included 16 from the first site and 15 from the second site. Five students from each site were ultimately selected to participate in the small group interviews. The researcher selected the interview participants based on demographic information to ensure a fair representation of the entire group. One of the ten selected students chose not to participate in the interviews, yielding a final count of five participants at Site One and four at Site Two. The interviews were recorded, transcribed, and analyzed through NVivo. All of the data collected was triangulated by the researcher for validity purposes and to provide a clearer understanding of the findings (Creswell & Guetterman, 2019).

The qualitative findings collected during the small group interviews enriched and clarified the quantitative results. The transcripts from the two, small group interview sessions were analyzed and coded for thematic recognition (Saldana, 2016). The researcher not only looked for the themes that would likely appear based on the theoretical framework but also investigated any such themes that were expected to appear and did not (Marshall & Rossman,

2016). Based on the theoretical framework for this study, the qualitative analysis found the major themes of learning environment, grading perceptions, self-efficacy, and parents. These findings were merged with the quantitative data collected in phase one and the theoretical framework of academic self-efficacy to develop a clear understanding of the participants perceptions (Marshall & Rossman, 2016).

Research Question #1: Summary of Results and Discussion

The first question the researcher sought to answer with the data was, "What is the impact of standards-based grading on a high school student's academic self-efficacy?" Both qualitative and quantitative data analysis from this study suggest that the participant's overall general self-efficacy score significantly improved with a standards-based grading practice when compared to their perceived self-efficacy in a traditional grading practice setting. This question was analyzed through the GSE survey that was administered to all participants. Participants were asked to complete the GSE based on their current perspective in a standards-based grading classroom and then answer the same instrument based on their experience in previous traditional graded classrooms.

Research has shown that academic self-efficacy is strongly related to a person's ability to perform and adjust in the post-secondary environment (Chemers et al., 2001; Honicke & Broadbent, 2015). Survey results indicated that the participant's GSE scores increased significantly after the implementation of a standards-based grading philosophy (Table 12). The overall combined GSE score saw a mean increase from 31 out of a possible 40 to 32.78 out of a possible 40 with a median increase from 31 to 33. Supporting evidence for this increase was also observed through the qualitative interviews under the theme of self-efficacy.

Table 16.

Rank by Time in SBG

Time in SBG	GSE Positive Rank	GSE Negative Rank	GSE Tie	Precent Positive
0-1 Year	1	2	0	33%
1-2 Years	10	4	7	71%
More than 2 years	24	10	5	71%

Note. Standards-based grading (SBG)

The data provided in Table 16 indicates that after the first year of implementation, 71% of the students had increased general self-efficacy scores. This suggests that, although the initial implementation of standards-based grading into a classroom can bring self-efficacy scores down, after the first-year students perceive higher efficacy scores overall.

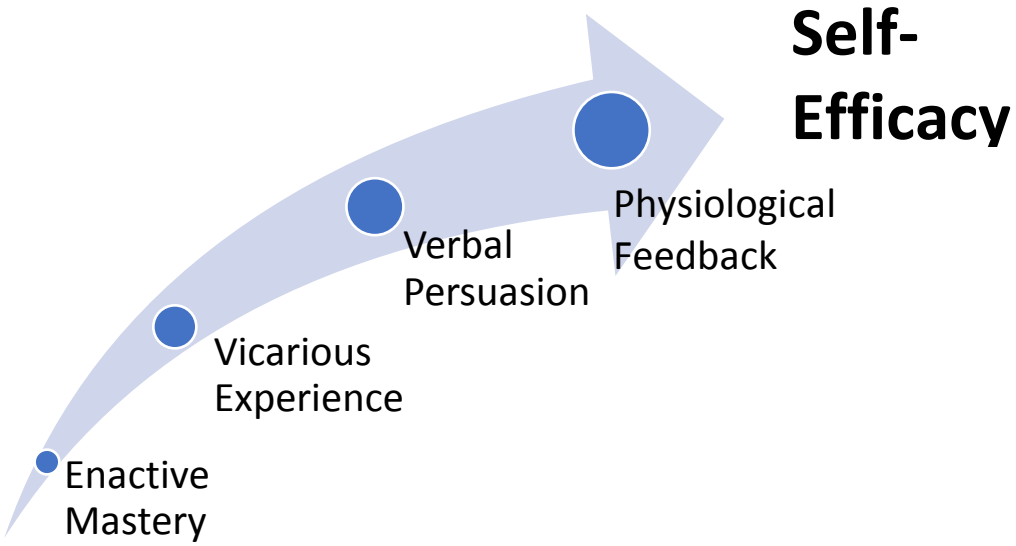
The survey results indicate that students working in a classroom using a standards-based grading philosophy have significantly improved self-efficacy scores. The finding of an increase in overall self-efficacy is especially significant with high school juniors because self-efficacy beliefs have been shown to decrease as students advance through school (Pintrich & Schunk, 1996; Schunk & Pajares, 2002). One of the factors attributed to this decrease in self-efficacy in school is norm-referenced or traditional grading (Pintrich & Schunk, 1996; Schunk & Pajares, 2002). Based on these results, the effective implementation of a standards-based grading program at the high school level will have a positive impact on a student's academic self-efficacy. Although concerns have been expressed from parents regarding the impact on post-secondary education (Peter & Buckmiller, 2014; Townsley, 2019), the findings in this study suggest that a standards-based grading philosophy increases self-efficacy which has been shown

to have a significant impact on academic performance and college outcomes (Chemers et al., 2001; Gore, 2006).

The findings in this study align with the four major influences on the development of a person's self-efficacy, as illustrated in figure 11:

Figure 11.

Influences on Self-Efficacy



Bandura 1986

The way someone is graded and given feedback on their academic performance has an impact on a person's self-efficacy and, in turn, their vicarious experience, enactive mastery, verbal persuasion, and physiological feedback (Pintrich & Schunk, 1996; Schunk & Pajares, 2002). This was evidenced through the results of the GSE data and confirmed through the qualitative interviews. Many of the comments documented during the semi-structured, small group interviews also confirmed an increased self-efficacy. All participants in the focus group interviews made statements indicating that the standards-based grading philosophy has had a

positive impact on a portion of the four areas identified that determine efficacy judgments. The following statements provide evidence of their perspectives:

- "…if you fail an assignment early in a traditional grading system, then you're probably in trouble for the rest of the quarter. Whereas in standards-based grading, you can always make that up, so there's the possibility of getting a higher grade."
- "...and that's one thing I do like about it is we do have the opportunity to go and fix that, and you know, get our grade to where we want, just by you know, if we have had a bad grade and it doesn't like, reflect on our final grade."
- "I think, like it's nice to have that chance to prove yourself because even, like I didn't like that some teachers wouldn't let us redo things because it like supposedly like doesn't show what we knew then, but I feel like, um, if you just have that clear understanding of, oh, I missed this but like I can still show that I understand it."
- "I honestly don't understand why… in the past, corrections and, like, redo's weren't available just because it's so prevalent, like sometimes there's miscommunication, there's confusion."
- "I like it a lot because my grades are better now and um, it's been really nice because I'll do it wrong, or sometimes I'll get some points off but then I'll talk to the teacher and it'll be like, "Oh you missed this and this," and I'll just add it in so that makes it good."
- "I do take more risks, I like, make it more my own style instead of like, if they don't like this, then I can always do this."
- "I tried my hardest to keep up and I was able to get good grades and I feel like if it was traditional, or the other form of grading where you were, there's no redos, no late work, that kind of stuff, I would have been falling behind and it wouldn't necessarily have

> been my fault, like it would have...but it would have been outside factors that play a part in it. I feel like in standards-based grading, the standard is the clear goal and I just need to get to that goal."

These statements indicate the positive impact that participants felt toward being able to demonstrate proficiency through multiple means without penalty. Additional comments in the interviews supported the idea that students were more willing to take on challenging work and less concerned about failure. Students explained that the chance to retake assessment or redo assignments helped them "challenge" themselves in the classroom without fear of damage to their overall grade.

Students that did not show increased self-efficacy could be influenced by a variety of external factors. Variables that were evidenced as aggravating or mitigating factors included parent support and teacher effectiveness in implementation. Research also provides evidence that academic self-efficacy goes down through high school years indicating that a smaller drop or maintenance of self-efficacy scores are indications of grading reform effectiveness.

Evidence supports the importance and influence of a strong self-efficacy, which has been linked to improved cognitive performance, perseverance, and task selection (Chemers et al., 2001). Based on this evidence, providing academic programs that promote or encourage increased self-efficacy should be a hallmark of an effective school. Student statements, such as "…if you don't do good learning all the stuff up to it, and then where you finally master it at the end is where it really pays off," indicate that standards-based grading is effective at impacting a student's beliefs and perceptions about grading and learning.

Research Question #2: Summary of Results and Discussion

The second research question allowed the researcher to analyze the individual questions on the GSE scale for each student. Rather than looking at the overall GSE score, the researcher looked at the change in rank for each question in an attempt to answer the research question: "What relationships can be drawn between variables associated with the high school student population of schools in a standards-based grading system compared to their previous experience in a traditional grading system?"

The results provided in Table 13 show that every question on the GSE survey had an increased mean score after the implementation of standards-based grading. Five of the questions had significant increases with $p < .05$. The five questions that showed a significant increase are as follows:

1) I am confident that I can deal effectively with unexpected events. ($p= .002$, $d =.39$)
2) Thanks to my resourcefulness I know how to handle unforeseen situations. ($p= .034$, $d=.38$)
3) I can remain calm when facing difficulties because I can rely on my coping abilities. ($p=.032$, $d=.28$)
4) When I am confronted with a problem, I can usually find several solutions. ($p=.019$, $d= .30$)
5) If I am in trouble, I can usually think of a solution. ($p=.023$, $d= .31$)

The effect size for each individual question was calculated, listed above, and determined to be medium. The collective data show that standards-based grading practices not only impact the overall GSE scores for the students, but it also has a positive effect on a student's ability to deal with unexpected events, overcome challenges, and find solutions when faced with a problem. A

student's self-efficacy has been directly correlated with their academic fortitude, ability to learn, and performance in the classroom (Pajares, 1996; Schunk, 1995). "Compared with students who doubt their learning capabilities, those who feel efficacious for learning or performing a task participate more readily, work harder, persist longer when encounter difficulties, and achieve at a higher level" (Schunk & Pajares, 2002, p. 2-3). The information from the individual questions not only confirms this concept but also provides insight into the separate constructs that together make-up a person's self-efficacy.

The five questions that showed significant increases indicate the participants felt increased levels of confidence, problem solving, and coping. Each of these components are foundational constructs to a healthy self-efficacy in an academic setting (Bandura, 1986; Pajares 1996). "Knowledge, skill and prior attainments are often poor predictors of subsequent attainments because the beliefs that individuals hold about their abilities and about the outcome of their efforts powerfully influence the ways in which they will behave" (Pajares, 1996, p. 543). The five specific questions that showed significant increases provide evidence that students in a standards-based grading classroom hold stronger beliefs about their abilities to overcome challenges, deal with stress, and problem solve effectively as compared to their previous experience in a traditional grading model.

The qualitative data obtained through semi-structured interviews with selected students confirmed students' beliefs in their abilities to deal with unexpected events, overcome challenges, and find solutions when faced with a problem. Statements, such as "…say I struggled at the beginning and then at the end I finally brought my grade up to an A, then that's where it pays off," provided substantial validation of the quantitative data that was obtained. Other statements the supported the quantitative findings are listed below:

- “I feel like if I am struggling with an assignment, I will try it and I know I can always correct it if I need to.”
- “I don’t think school is as challenging, I mean I still learn the same but I don’t feel like everything needs to be perfect.”
- “I am more willing to take more risks with my work now.”
- “We do have the opportunity to go and fix that, and you know, get our grade to where we want, just by you know, if we have had a bad grade and it doesn’t like, reflect on our final grade.”
- “Like, it's nice to have that chance to prove yourself because even, like, I didn't like that some teachers wouldn't let us redo things because it like supposedly like doesn't show what we knew then, but I feel like, um, if you just have that clear understanding of, ‘Oh, I missed this but like I can still show that I understand it.’”

In summation, both the quantitative and qualitative data collected indicate that students in a standards-based grading system have increased perceptions of their ability to problem solve, deal with unexpected change, and overcome challenges in the classroom.

Conclusions

Grading academic progress is a foundational part of the American educational system in which outdated factors continue to misrepresent student achievements (Brookhart, 2009; Guskey, 2015; Townsley & Buckmiller, 2020). As schools look toward continual improvement, grading practices must be addressed to ensure validity and reliability in reporting (Brookhart, 2009; Guskey, 2015; Townsley & Buckmiller, 2020). In response to the need for a more accurate grading system, standards-based grading has been implemented at schools across the United States (Brookhart, 2009; Guskey, 2015; Townsley & Buckmiller, 2020). The implementation of

this new grading philosophy has presented challenges for those involved (Knight & Cooper, 2019; Schimmer, 2016). Specifically, the change initially creates more work for teachers (Diegelman-Parente, 2011) and parents have expressed concerns about post-secondary implications (Peters & Buckmiller, 2014; Schimmer 2016). It is difficult for school leadership to maintain a balance between reforming grading systems, as recommended by many scholars, and the foundational beliefs in traditional grading practices that are found in many stakeholders (Knight & Cooper, 2019). This research sought to provide empirical evidence focused on the impact that a standards-based grading platform has on the academic self-efficacy of students.

Self-efficacy is the belief or perceptions that a person has in their ability to learn or perform a task at a specific level, and it has been closely correlated with a student's academic performance (Bandura, 1997; Pajares, 2002; Usher & Pajares, 2006). In the academic realm, this self-efficacy has been associated with motivation, learning, task choice, and achievement (Pajares 1996; Schunk, 1995). It is critical to note that this study focused on the academic self-efficacy of students in their junior year of high school, a time when research has shown self-efficacy beliefs tend to decrease as students move through school (Pintrich & Schunk, 1996). Several school practices tend to lower a students' academic self-efficacy, especially for students who are not academically prepared for the challenges of school (Schunk & Pajares, 2002). Research has shown a strong link between academic self-efficacy and college performance, college persistence, and perceived career options (Lent, Brown, & Larkin, 1986; Gore, 2006).

The focus of this study was to determine the impact of standards-based grading practices on the academic self-efficacy of high school students. The level of self-efficacy for each student in the study was measured using the GSE scale. The quantitative results showed that students' overall self-efficacy scores, as measured on the GSE, were significantly higher when taking part

in a standards-based grading system. This information was used in conjunction with the qualitative data that provided further support of increased self-efficacy. The quantitative and qualitative data from this study indicate that students involved in a standards-based grading system had increased self-efficacy. Student statements, such as "I do take more risks" and "I feel less worried about my grade and more concerned about my understanding," provide valuable insight into how students view these grading practices. This is of particular interest as schools look to counteract the practices that have been identified for lowering a students' self-efficacy. Gaining the support of the community and helping those involved to understand the reasoning behind a standards-based grading philosophy is one of the most difficult obstacles educational leaders must overcome when implementing this new system (Guskey & Jung 2013; Peters et al., 2017; Spencer, 2012). This study provides empirical qualitative and quantitative evidence that indicates students in a standards-based grading system have increased levels of academic self-efficacy.

The second part of this study sought to take a deeper look at the individual questions on the GSE scale to determine how standards-based grading influenced the scores. The analysis showed all the questions had increased mean scores, with five of the ten questions on the GSE having statistically significant increases in scores. The effect size of these increases was all determined to be medium. The five questions related to a standards-based grading system that saw significant increases were centered around being able to overcome challenges, coping with change, and effectively problem solving. The qualitative data was designed to capture student voice and their perceptions of the standards-based grading system. Research supports the use of student voice as an effective mechanism to address educational issues (Mitra, 2004, 2008). The small group interviews provided clarity as to how students felt about standards-based grading,

the challenges they encountered, and the benefits they perceived. The themes derived from the interviews, combined with the qualitative data, clearly illustrated the correlation between this new grading system and a student's increased confidence, ability to problem solve effectively, and ability to deal with obstacles. These critical skills, as a part of overall academic self-efficacy, have been correlated to post-secondary success and career options after high school (Chemers et al., 2001; Gore, 2006).

Recommendations for Further Research

The findings from this mixed methods study provide insight into high school student perceptions of grading practices. This mixed methods study was designed to provide educational entities evidence of the impact that standards-based grading has on academic self-efficacy. The findings in this study provide stakeholders with some insight into how school students perceive the changes in grading philosophies and evidential reasoning to support a change. Further research will be critical in developing patterns of success and challenges that are presented as schools look to reform grading practices. Therefore, the first recommendation for further research is to expand the sample size involved in the study. Replicating the study with larger groups and/or at more locations would give a more generalizable depiction of the impact of standards-based grading on a student's self-efficacy.

The second recommendation also requires a larger sample size, but one that also focuses on a wider demographic range. Due to the lack of diversity in the sample population, a larger, more diverse, sample size would help in determining the grading system's impact on different groups of students. Incorporating a larger sample size should also include replicating this study with a younger population. Looking at the impact of standards-based grading practices on elementary school students is a key component that needs further investigation.

In addition, replicating this research with younger students will provide valuable insight into age specific impacts of standards-based grading implementation. Conducting research in middle schools and elementary schools would provide data that give insight into the best age groups to transition grading practices. It would also allow for a comparative analysis of the impacts of grading reform implementation in high school versus middle school versus elementary school students.

Another recommendation is for research on students in a standards-based grading system that has been implemented for a longer period of time. Both schools involved in this study had been using standards-based grading practices for fewer than five years. Involving students that attend schools with longer held standards-based grading practices would provide a perspective on the long-term impacts on students. This additional research would provide further insight into existing research on student's perceptions of standards-based grading practices (Peters et al., 2017).

As more schools work to reform their grading practices, it is critical that research continues to investigate the correlation between standards-based grading practices and standardized tests. Although academic self-efficacy has been strongly correlated with academic performance, connecting standards-based grading to improved test scores is crucial as educators look to gain support from all stakeholders.

Implications for Professional Practice

The results of this study provide stakeholders with valuable data regarding student perceptions of standards-based grading. This should be reviewed by anyone interested in implementing new grading practices as they gain insight into student beliefs about grading reform. Obtaining student voice in educational initiatives has been shown to improve outcomes

(Mitra, 2008). It is still imperative to include students in grading conversations within individual schools, but this research provides some insight into student perceptions that can be expected throughout this transition. The research provided in this study provides all stakeholders with valuable information regarding the impact of the work that is being done. This research demonstrates the power that grading reform can have on student self-efficacy and can provide rationale a reason for resiliency as institutions face the expected challenges that are inherent with transforming grading practices.

Although student self-efficacy is not commonly discussed at the building level, such constructs as motivation, academic performance, perseverance, and problem-solving abilities are always a topic for educators. Self-efficacy has been shown to be a powerful predictor of academic performance and post-secondary perseverance (Gore, 2006; Chemers et al., 2001). This study links these attributes associated with self-efficacy and grading practices. This is a critical step in providing evidential support of standards-based grading practices. These results should be used to illustrate the reasoning and rationale behind making substantial adjustments in current grading practices. These results also provide current educational leaders and scholars with a clear connection between standards-based grading and academic performance. This research provides the "why" behind the change in grading practices with both qualitative and quantitative evidence to support the change in practices.

For administrators at the building and district level, this data provides additional evidence to support the transition into a standards-based grading system. The benefits of accurately reporting student progress through the effective implantation of standards-based grading is widely known (Brookhart, 2009; Guskey, 2015; Schimmer, 2016), but this research shows the positive impact that this practice has on a high school student's self-efficacy. This is critical to

get in the hands of stakeholders that push back and try to hold onto traditional grading practices. Additionally, it provides a measuring device to determine the level of effective grading reform. Looking at academic self-efficacy through the GSE throughout a grading reform process will provide leaders with a pulse on students as they are working through this change.

As part of any educational initiative, teacher voice becomes a crucial component (Quaglia & Lande, 2016). This research should be used to start the conversation about the importance of reflecting on our grading practices. As a part of this reflection, educating teachers on the impact of self-efficacy on student achievement is critical. Transitioning that conversation into a reflection on systems that promote self-efficacy provides a solid foundation to begin imperative grading conversations. For teachers looking at standards-based grading, this study provides insight and perspective into how grading practices can determine the efficaciousness of a student in class. As conversations around grading begin to happen on campuses, this provides further evidence that the current practices are not only ineffective, but that they are also harmful. Teachers report that the initial implementation of standards-based grading creates more work (Hill, 2018), which makes it crucial for them to understand the impact of these grading practices (Townsley & Buckmiller, 2020). This information is pivotal as teachers explore the "why" behind adjusting their long-held grading perceptions.

Parents and students can be the most obstinate group of stakeholders needing to be brought on board as teammates when educational institutions look to make transitions in grading practices (Schimmer, 2016). As parents express concerns about post-secondary implications or lack of understanding regarding this new grading system begins to surface (Franklin et al., 2016; Reide, 2018; Yost, 2015), this study provides valuable information to help parents understand the positive impact that changes in the grading system will have on students. Student voice, which is

documented throughout this study, is a vital part of the process as schools look to make this transition. Knowledge and awareness of the powerful, positive impact that a standards-based grading system has on students is a key element in the effective implementation of such practices.

In summary, this study provides all stakeholders with the big picture impacts of standards-based grading in three or less years. The sites use for the study saw significant increases in academic self-efficacy shortly after the implementation of standards-based grading practices. This information should be used to guide, support, and inform stakeholders throughout the implementation process. The data definitively supports the concept that standards-based grading practices either increase or maintain academic self-efficacy levels in high school students. This should be communicated to parents, teachers, and students in an effort to garner support during the transition of grading philosophies. Starting the conversation about academic self-efficacy of students rather than grading can be a powerful first step in engaging stakeholders that are resistant to grading reform. Additionally, the qualitative evidence obtained, provides guidance from students regarding the pitfalls of the implementation of standards-based grading practices. Reviewing this information will help school leaders proactively develop systems to address concerns such as clear parent communication. This research should be critically analyzed and used by all stakeholders to create a strategic implementation plan prior to starting any grading reform.

Lastly, the use of self-efficacy to measure the effectiveness of systems in place is something that can be transferred into all areas of education. This provides another data point outside of standardized tests and grades that help determine the effectiveness of a system. The research model used in this study can easily be transferred from research on standards-based

grading into any other educational reform movement. School leaders using academic self-efficacy rates as a gauge to determine the effectiveness of a system will provide clarity regarding the implementation of a variety of educational initiatives.

References

Airasian, P. W. (1997). *Classroom assessment*. New York, NY: McGraw Hill.

Allen, J. D. (2005). Grades as valid measures of academic achievement of classroom learning. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas, 78*(5), 218-223.

Anderson, L. W. (2018). A critique of grading: Policies, practices, and technical matters. *Education Policy Analysis Archives, 26*(49), 1-31.

Bailey, J., & McTighe, J. (1996). Reporting achievement at the secondary level: What and how. *Association for Supervision and Curriculum Development-Yearbook*, 119-140.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychology Review, 84*(2), 191-215.

Bandura, A. (1986). *Social foundations of thought and action*. Englewood Cliffs, NJ: Prentice Hall

Bandura, A. (1991). Social cognitive theory of self-regulation. *Organizational Behavior and Human Decision Processes, 50*(2), 248-287.

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist, 28*(2), 117-148.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Bandura, A., Claudio, B., Caprara, G. V., & Pastoreli, C. (1996). Multifaceted impact of self-efficacy beliefs on academic functioning. *Child Development, 67*(3), 1206–1222. https://doi.org/10.2307/1131888

Bandura, A., & Locke, E. A. (2003). Negative self-efficacy and goal effects revisited. *Journal of Applied Psychology, 88*(1), 87-99.

Bandura, A., & Schunk, D.H. (1981). Cultivating competence, self-efficacy, and intrinsic interest through proximal self-motivation. *Journal of Personality and Social Psychology, 41*(3), 586-598.

Beatty, I. D. (2013). Standards-based grading in introductory university physics. *Journal of the Scholarship of Teaching and Learning*, *13*(2), 1-22.

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

Brookhart, S. M. (2009). *Grading (2nd ed.).* New York, NY: Merrill Pearson Education.

Brookhart, S.M. (2011). Starting the conversations about grading. *Educational Leadership, 69*(3), 10-14. Retrieved from https://www.greatschoolspartnership.org/wp-content/uploads/2016/11/Starting-the-Conversation-about-Grading-2.pdf

Brookhart, S.M., Guskey, T.R., Bowers, A.J., McMillan, J.H., Smith, J.K., Smith, L.F., & Welsh, M.E. (2016). A century of grading research: Meaning and value in the most common educational measure. *Review of Education Research, 86*(4), 803-848 doi: 10.3102/0034654316672069

Campbell, A.L. (1921). Keeping the score. *School Review*, *29*(7), 510-519.

Chemers, M.M., Garcia, B.F., & Hu, L. (2001). Academic self-efficacy and first-year college student performance and adjustment. *Journal of Educational Psychology, 93*(1), 55-64. doi:10.1037//0022-0663.93.1.55

Chory-Assad, R. M. (2002). Classroom justice: Perceptions of fairness as a predictor of student motivation, learning, and aggression. *Communication Quarterly*, *50*(1), 58–77. https://doi.org/10.1080/01463370209385646

Chory, R. M., & McCroskey, J. C. (1999). The relationship between teacher management communication style and effective learning. *Communication Quarterly, 47*, 1–11. https://doi.org/10.1080/01463379909370120

Colosi, L., & Dunifon, R. (2006). What's the difference? "Post then pre" & "pre then post". *Cornell Cooperative Extension.* Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.91.6252&rep=rep1&type=pdf

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

Creswell, J. W. (2015). *30 essential skills for the qualitative researcher*. Sage Publications.

Creswell, J. W., & Guetterman, T. C. (2019). *Educational research: Planning, conducting and evaluating quantitative and qualitative research (Sixth Edition).* New York: Pearson.

Davis, G. A. (2003). Using a retrospective pre-post questionnaire to determine program impact.

Davis, L. (2019, February 13). Standards-based grading: What to know in 2019. *Schoology Exchange.* https://www.schoology.com/blog/standards-based-grading

Davis, L. (2020, January 13). Standards-based grading: What to know in 2020. *Schoology Exchange.* https://www.schoology.com/blog/standards-based-grading

Di Giunta, L., Alessandri, G., Gerbino, M., Kanacri, P. L., Zuffiano, A., & Caprara, G. V. (2013). The determinants of scholastic achievement: The contribution of personality traits, self-esteem, and academic self-efficacy. *Learning and individual Differences*, *27*, 102-108.

Diegelman-Parente, A. (2011). The use of mastery learning with competency-based grading in an organic chemistry course. *Journal of College Science Teaching*, *40*(5).

Duckworth, A. (2016). *Grit: The Power and Passion of Perseverance.* Simon & Schuster.

Dweck, C.S. (1986). Motivational processes affecting learning. *American Psychologist, 41*(10), 1040-1048. https://doi.org/10.1037/0003-066x.41.10.1040

Dweck, C. S. (2008). *Mindset: The new psychology of success*. Random House Digital, Inc..

Espenshade, T. J, Lynch, S. M., &. Zajacova, A. (2005). Self-efficacy, stress, and academic success in college. *Research in Higher Education*, *46*(6), 677–706. https://doi.org/10.1007/s11162-004-4139-z

Fay, C. H., & Gordon, M. E. (2010). The effects of grading and teaching practices on students' perceptions of grading fairness. *College Teaching*, *58*(3), 93-98.

Field, A. (2018). *Discovering statistics using IBM SPSS statistics.* Sage.

Fielding, M. (2001, June). Beyond the rhetoric of student voice: New departures or new constraints in the transformation of 21st century schooling?. In *Forum for promoting 3-19 comprehensive education* Vol. 43, No. 2, pp. 100-109.

Finkelstein, I. E. (1913). *The marking system in theory and practice* (No. 10). Baltimore: Warwick & York.

Franklin, A., Buckmiller, T., & Kruse, J., (2016). Vocal and vehement: Understanding Parents' aversion to standards-based grading. *International Journal of Social Science Studies*, *4*(11), 19–29. doi.10.11114/ijsss.v4i11.1923

Frisbie, D. A., & Waltman, K. K. (1992). Developing a personal grading plan. *Educational Measurement: Issues and Practice*, *11*(3), 35-42.

Frisbie, D. A., & Waltman, K. K. (1994). Parents' understanding of their children's report card grades. *Applied Measurement in Education*, *7*(3), 223-240.

Gladwell, M. (2008). *Outliers: The story of success*. Hachette UK.

Gore, P. A., Jr. (2006). Academic self-efficacy as a predictor of college outcomes: Two incremental validity studies. *Journal of career assessment, 14*(1), 92-115.

Guskey, T. R. (1996). Communicating Student Learning. 1996 ASCD Yearbook.

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

Guskey, T. R. (2011). Five obstacles to grading reform. *Educational Leadership, 69*(3), 16-21. Retrieved from https://uknowledge.uky.edu/cgi/viewcontent.cgi?article=1006&context=edp_facpub

Guskey, T. R. (2013). The case against percentage grades. *Educational Leadership, 71*(1), 68-72.

Guskey, T.R. (2015). *On your mark: Challenging the conventions of grading and reporting.* Solution Tree Press.

Guskey, T. R., & Bailey, J. M. (2001). *Developing grading and reporting systems for student learning.* Corwin Press.

Guskey, T.R, &. Bailey, J.M. (2010). *Developing standards-based report cards.* Thousand Oaks, CA: Corwin.

Guskey, T. R., & Jung, L. A. (2012). Four Steps in Grading Reform. *Principal Leadership*, *13*(4), 22-28.

Guskey, T. R., & Jung, L.A. (2013). *Answers to essential questions about standards, assessments, grading, and reporting.* Corwin Press.

Guskey, T. R., & Link, L.J. (2019). Exploring the factors teachers consider in determining students' grades. *Assessment in Education: Principles, Policy & Practice*, *26*(3), 303-320. doi:10.1080/0969594X.2018.1555515

Guskey, T.R., & Muñoz, M. A. (2015). Standards-based grading and reporting will improve education. *Phi Delta Kappan, 96*(7), 64–68. doi:10.1177/0031721715579043.

Gutek, G. L. (1986). *Education and the United States: A historical perspective.* Englewood Cliffs, NJ: Prentice Hall.

Hannon, B. A. (2014). Predicting college success: The relative contributions of five social/personality factors, five cognitive/learning factors, and SAT scores. *Journal of Education and Training Studies, 2,* 46-58. doi:10.11114/jets.v2i4.451

Hattie, J. (2012). *Visible learning for teachers maximizing impact on learning*. New York, NY: Routledge.

Heflebower, T., & Hoegh, J. K. (2014). *A school leader's guide to standards-based grading*. Solution Tree Press.

Heflebower, T., Hoegh, J. K., Warrick, P. B., Hoback, M., McInteer, M., Clemens, B., & Marzano, R. J. (2014). *A school leader's guide to standards-based grading*. Bloomington, IN: Marzano Research Laboratory.

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

Hill, G. R. (2018). *Impact of teacher attitudes on implementation of a standards-based grading system* (Doctoral dissertation). Retrieved from ProQuest. (ProQuest No. 13426348)

Honicke, T., & Broadbent, J. (2015). The influence of self-efficacy on academic performance: A systematic review. *Educational Research Review, 17*(1), 63-84. http://dx.doi.org/10.1016/j.edurev.2015.11.002

Iamarino, D. L. (2014). The benefit of standards-based grading: A critical evaluation of modern grading practices. *Current Issues in Education, 17*(2), 1-12. Retrieved from http://cie.asu.edu/ojs/index.php/cieatasu/article/view/1234

Kirschenbaum, H., Simon, S.B., & Napier, R. W. (1971). *Wad-ja-get? The grading game in American education.* New York: Hart.

Knight, M., & Cooper, R. (2019). Taking on a new grading system: The interconnected effects of standards-based grading on teaching, learning, assessment, and student behavior. *National Association of Secondary School Principals Publication, 103*(1) 65-92. https://doi.org/10.1177/019263650108562407

Kohn, A. (1999). From degrading to de-grading. *High school magazine*, *6*(5), 38-43.

Kohn, A. (2011). The case against grades. *Educational Leadership, 69*(11), 28-33.

Komarraju, M., & Nadler, D. (2013). Self-efficacy and academic achievement: Why do implicit beliefs, goals, and effort regulation matter? *Learning and Individual Differences*, *25*, 67-72.

Kunnath, J. P. (2017). Teacher grading decisions: Influences, rationale, and practices. *American Secondary Education, 3*, 68-88.

Lee, E., Carberry, A., Diefes-Dux, H., Atwood, S., & Siniawski, M. (2017). Faculty perception before and after implementation of standard-based grading. In *Research in Engineering Education Symposium* (pp. 6-8).

Lent, R. W., Brown, S. D., & Larkin, K. C. (1986). Self-efficacy in the prediction of academic performance and perceived career options. *Journal of Counseling Psychology, 33*(3), 265-269.

Marshall, C., & Rossman, G. B. (2014). *Designing qualitative research.* Sage publications.

Marzano, R. (2000). *Transforming classroom grading*. Alexandria, VA: ASCD.

Marzano, R. J. (2010). *Formative assessment & standards-based grading*. Bloomington, IN: Marzano Research.

Marzano, R. J., & Boogren, T. (2010). *Developing expert teachers*. Solution Tree Press.

Maxwell, J. A. (2012). *A realist approach for qualitative research.* Sage.

Mitra, D. L. (2004). The significance of students: can increasing "student voice" in schools lead to gains in youth development? *Teachers College Record, 106*, 651-688.

Mitra, D. (2008). Amplifying student voice. *Educational Leadership, 66*(3).

Nitko, A. J., & Brookhart, S. M. (2007). Education assessment of students. *New Jearsey: Merrill Prentice Hall*.National Center for Educational Statistics. (2018). Fast facts: Back to school statistics. Retrieved May 21, 2019, from https://nces.ed.gov/fastfacts/display.asp?id=372

O'Connor, K. (2002). *How to grade for learning: Linking grades to standards* (2nd ed.). Thousand Oaks, CA: Corwin.

O'Connor, K. (2007). *A repair kit for grading.* Portland, OR: Educational Testing Service.

O'Connor, K. (2018). *How to grade for learning.* Thousand Oaks, CA: Corwin.

O'Connor, K., Jung, L.A., & Reeves, D. (2018). Gearing up fast grading and reporting: It is time schools move toeard a grading system that is fair, accurate, specific, and timely. *Phi Delta Kappan, 99*(8), 67-71. https://doi.org/10.1177/0031721718775683

Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research, 66*(4), 543-578.

Pajares, F. (2002). Self-efficacy beliefs in academic contexts: An outline. https://www.uky.edu/~eushe2/Pajares/efftalk.html

Pajares, F. (2009). Toward a positive psychology of academic motivation. In R. Gilman, E. S. Huebner, & M. J. Furlong (Eds.), *Handbook of positive psychology in schools* (p. 149–160). Routledge/Taylor & Francis Group.

Pajares, F., & Schunk, D. H. (2001). Self-beliefs and school success: Self-efficacy, self-concept, and school achievement. *Perception, 11*(2), 239-266.

Pajares, F., & Usher, E. L. (2006). Inviting confidence in school: Invitations as critical sources of the academic self-efficacy beliefs of entering middle school students. *Journal of Invitational Theory and Practice, 12*, 7-16.

Peters, E. (2009). Minimum grades no longer allowed. *Abilene Reporter-News*. Retrieved from www.reporternews.com/news/2009/oct/20/ minimum-grades

Peters, R., & Buckmiller, T. (2014). Our grades were broken: Overcoming barriers and challenges to implementing standards-based grading. *Journal of Educational Leadership in Action*, *2*(2), 1-23.

Peters, R., & Buckmiller, T. (2015). Our grades were broken: Overcoming barriers and challenges to implementing standards-based grading. *Journal of Educational leadership in Action 2*(2). Retrieved from: http://www.mrshutt.org/uploads/4/6/3/8/46385679/our_grades_were_broken__peters_and_buckmiller_.pdf

Peters, R., Kruse, J., Buckmiller T., & Townsley, M. (2017). "It's just not fair!" Making sense of secondary students' resistance to a standards-based grading. *American Secondary Education*, *45*, 9–28. Retrieved from https://eric.ed.gov/?id=EJ1154237

Pintrich, R. P., & Schunk, D. H. (1996). Motivation in education. *Theory, Research, and Application.* New Jersey: Prentice Hall.

Pollio, M., & Hochbein, C. (2015). The association between standards-based grading and standardized test scores in a high school reform model. *Teachers College Record, 117*(11).

Quaglia, R. J., & Lande, L. L. (2016). *Teacher voice: Amplifying success.* Corwin Press.

Redmond, B., last modified by McNabb, H. (2010-2016). PSYCH 484 Wiki: 7. Self-Efficacy and Social Cognitive Theories. Retrieved on February 27, 2016 from: 7. Self-Efficacy and Social Cognitive Theories

Reeves, D. B. (Ed.). (2008). *Assessing educational leaders: Evaluating performance for improved individual and organizational results*. Corwin Press.

Richmond, E. (2008, February 4). *A floor for failing grades: Parents, educators debate whether kids should get 50 points for doing nothing.* Las Vegas Sun.

Riede, P. (2018). Making the call inside admissions offices. *School Administrator, 75*(2), 26-29.

Ripski, M. B., & Gregory, A. (2009). Unfair, unsafe, and unwelcome: Do high school students' perceptions of unfairness, hostility, and victimization in school predict engagement and achievement?. *Journal of School Violence*, *8*(4), 355-375.

Rockwell, S. K., & Kohn, H. (1989). Post-then-pre evaluation. *Journal of Extension*, *27*(2), 19-21.

Rudduck, J., & Flutter, J. (2000). Pupil participation and perspective: 'Craving a new order of experience'. *Cambridge Journal of Education, 30*(1), 75-89.

Saldaña, J. (2016). The coding manual for qualitative researchers. SAGE Publications Limited.

Scarlett, M. (2018). "Why did I get a C?": Communicating student performance using standards-based grading. *A Journal of Scholarly Teaching, 13*, 59-75.

Schimmer, T. (2016). *Grading from the inside out*. Bloomington, IN: Solution Tree Press

Schunk, D.H. (1991). Self-efficacy and academic motivation. *Educational Psychologist, 26*(3-4), 207-231. https://doi.org/10.1080/00461520.1991.9653133

Schunk, D. H. (1995). Self-efficacy and education and instruction. In *Self-efficacy, adaptation, and adjustment* pp. 281-303. Boston, MA: Springer

Schunk, D. H., & Pajares, F. (2002). The development of academic self-efficacy. In *Development of achievement motivation* (pp. 15-31). Academic Press.

Schwarzer, R., & Jerusalem, M. (1995). Optimistic self-beliefs as a resource factor in coping with stress. In *Extreme stress and communities: Impact and Intervention* (pp. 159-177). Dordrecht: Springer.

Scriffiny, P.L. (2008). Seven Reasons for Standards-Based Grading. Educational Leadership, 66(2), 70–74. Retrieved from https://search-ebscohost-com.ezproxy.hamline.edu/login.aspx?direct=true&db=e ft&AN=508004154&site=ehost-live.

Spencer, K. (2012). Standards-based grading. *The Education Digest*, *78*(3), 4-10.

Starch, D., & Elliot, E.C. (1912). Reliability of the grading of high school work in English. *School Review*, *20*, 442-457.

Starch, D., & Elliot, E.C. (1913). Reliability of the grading of high school work in mathematics. *School Review*, *21*, 254-259.

Swan G. M., Guskey, T. R., & Jung, L. A. (2014). Parents' and teachers' perceptions of standards based and traditional report cards. *Education Assessment, Evaluation and Accountability, 23*, 289-299. doi: 10.1007/s11092-014-9191-4

Tierney, R. D., Simon, M., & Charland, J. (2011). Being fair: Teachers' interpretations of principles for standards-based grading. *The Educational Forum*, *75*(3), 210-227.

Townsley, M. (2019). Considering standards-based grading: Challenges for secondary school leaders. *Journal of School Administration Research and Development, 4*(2), 35-38. Retrieved from https://eric.ed.gov/?id=EJ1228578

Townsley, M., & Buckmiller, T. (2016, January). What does the research say about standards-based grading? [Blog post]. Retrieved from http://mctownsley.net/standards-based-grading- research/

Townsley, M., & Buckmiller, T. (2020). Losing As and Fs: What works for schools implementing standards-based grading? *Educational Considerations*, *46*(1), 1-7. Retrieved from https://newprairepress.org/edconsderations/vol46/iss1/3

Townsley, M., & Varga, M. (2018). Getting high school students ready for college: A quantitative study of standards-based grading practices. *Journal of Research in Education, 28*(1), 92–112. Retrieved from https://files.eric.ed.gov/fulltext/EJ1168171.pdf

Vatterott, C. (2015). *Rethinking grading: Meaningful assessment for standards-based learning.* ASCD.

Welsh, M. E., D'Agostino, J. V., & Kaniskan, B. (2013). Grading as a reform effort: Do standards-based grades converge with test scores? *Educational Measurement: Issues & Practice, 32*(2), 26–36. https://doi.org/10.1111/emip.12009

Wormeli, R. (2011). Redos and retakes done right. *Educational Leadership, 69*(3), 22-26.

Yost, R. (2015). Standards-based proposal draws concerns in Forest city. *The Globe Gazette.* Retrieved from http://globegazette.com/news/local/standards-based-proposal-drawsconcerns-in-forest-city/article_c3c640fd-bade-5d83-b353-c6a4a235653f.html

Appendix A

General Self Efficacy Scale

General Self-Efficacy Scale (GSE)

About: This scale is a self-report measure of self-efficacy.

Items: 10

Reliability: Internal reliability for GSE = Cronbach's alphas between .76 and .90

Validity: The General Self-Efficacy Scale is correlated to emotion, optimism, work satisfaction. Negative coefficients were found for depression, stress, health complaints, burnout, and anxiety.

Scoring: Not at all true
Hardly true Moderately true
Exactly true
All questions 1 2 3 4

The total score is calculated by finding the sum of all the items. For the GSE, the total score ranges between 10 and 40, with a higher score indicating more self-efficacy.

References: Schwarzer, R., & Jerusalem, M. (1995). Generalized Self-Efficacy scale. In J. Weinman, S. Wright, & M. Johnston, Measures in health psychology: A user's portfolio. Causal and control beliefs (pp. 35-37). Windsor, UK: NFER-NELSON.

General Self-Efficacy Scale (GSE)

	Not at all true	Hardly true	Moderately true	Exactly true
1. I can always manage to solve difficult problems if I try hard enough	☐	☐	☐	☐
2. If someone opposes me, I can find the means and ways to get what I want.	☐	☐	☐	☐
3. It is easy for me to stick to my aims and accomplish my goals.	☐	☐	☐	☐
4. I am confident that I could deal efficiently with unexpected events.	☐	☐	☐	☐
5. Thanks to my resourcefulness, I know how to handle unforeseen situations.	☐	☐	☐	☐
6. I can solve most problems if I invest the necessary effort.	☐	☐	☐	☐
7. I can remain calm when facing difficulties because I can rely on my coping abilities.	☐	☐	☐	☐
8. When I am confronted with a problem, I can usually find several solutions.	☐	☐	☐	☐
9. If I am in trouble, I can usually think of a solution	☐	☐	☐	☐
10. I can usually handle whatever comes my way.	☐	☐	☐	☐

Appendix B

Adapted Survey

Q1 Hello. Thank you for taking the time to take the following survey. I am seeking information about how standards-referenced grading in your classroom, did or did not impact your perception of how you attack challenging tasks. You have the option to take this survey, but I hope you will participate. The results of the survey will help provide some insight educational leaders both in Idaho and around the world. Once the study is complete, I will destroy the survey. Your name and the name of the school will not appear in any reports and I will keep the information confidential. Thank you for your time.

☐ I agree to voluntarily participate in this survey

☐ I do not agree to participate in this survey

Q2 What is your current grade level?

☐ 11^{th}

☐ 12^{th}

Q3 I am a:

☐ Male

☐ Female

☐ I do not wish to self-identify

Q4 Please specify your ethnicity:

☐ White

☐ Hispanic or Latino

☐ Black or African American

☐ Native American or American Indian

☐ Asian / Pacific Islander

☐ Other

☐ I do not wish to self-identify

Q5 My mother has completed a high school diploma.

☐ Yes

☐ No

☐ Not sure

Q6 My father has completed a high school diploma.
☐ Yes
☐ No
☐ Not sure

Q7 My mother has finished a college degree.
☐ Yes
☐ No
☐ Not sure

Q8 My father has finished a college degree.
☐ Yes
☐ No
☐ Not sure

Q9 For each of the following statements, please indicate how standards-referenced grading has CHANGED your thoughts. Answer on the scale from Strongly Agree to Strongly Disagree.

	Before Standards Referenced Grading				After Standards Referenced Grading			
	Strongly Agree	Agree	Disagree	Strongly Disagree	Strongly Agree	Agree	Disagree	Strongly Disagree
I can always manage to solve difficult problems if I try hard enough.	☐	☐	☐	☐	☐	☐	☐	☐
If someone opposes me, I can find the means and ways to get what I want.	☐	☐	☐	☐	☐	☐	☐	☐
It is easy for me to stick to my aims and accomplish my goals.	☐	☐	☐	☐	☐	☐	☐	☐

I am confident that I can deal efficiently with unexpected events.	☐	☐	☐	☐	☐	☐	☐	☐
When I get a bad grade on an assignment, I feel deflated.	☐	☐	☐	☐	☐	☐	☐	☐
I can solve most problems if I invest the necessary effort.	☐	☐	☐	☐	☐	☐	☐	☐
I like an assignment best when I can do it perfectly the first time.	☐	☐	☐	☐	☐	☐	☐	☐
When I am confronted with a problem, I usually find several solutions.	☐	☐	☐	☐	☐	☐	☐	☐
If I am in trouble, I can usually think of a solution.	☐	☐	☐	☐	☐	☐	☐	☐
I can usually handle whatever comes my way.	☐	☐	☐	☐	☐	☐	☐	☐

Appendix C

Informed Consent Form

Consent for Minor Participation

A. PURPOSE AND BACKGROUND

Derek Bub, a doctoral student in the Department of **education** at Northwest Nazarene University is conducting a research study related to **standards-referenced grading**.

You are being asked to give consent for your child to participate in this study because **we would like to analyze the impact of standards-based grading on the academic self-efficacy of students.** Their participation will help researchers **understand student perspectives regarding different means of communicating academic progress.**

B. PROCEDURES

If you agree to be in the study, the following will occur:

1. You will be asked to sign an Informed Consent Form giving permission for your child to participate in this study.
2. Your child will be asked to participate in the **general self-efficacy and mindset survey** no more than once this semester.
3. Your child may be asked to participate in a short focus group with a researcher and their peers. In this focus group they will be asked to answer a set of focus group questions and engage in a discussion on **standards-based grading**. This discussion will be audio taped and is expected to last approximately 45-60 minutes.

These procedures will be competed at a location mutually decided upon by the participant and researcher and will take a total time of about 15 minutes.

C. RISKS/DISCOMFORTS

1. Some of the discussion questions may make your child uncomfortable or upset, but they are free to decline to answer any questions they do not wish to answer or to stop participation at any time.

2. For this research project, the researchers are requesting demographic information. Due to the make-up of Idaho's population, the combined answers to these questions may make an individual person identifiable. The researchers will make every effort to protect confidentiality. However, if you are uncomfortable answering any of these questions, your child may decline to answer them.

3. Confidentiality: Participation in research may involve a loss of privacy; however, your records will be handled as confidentially as possible. No individual identities will be used in any reports or publications that may result from this study. All data from notes, audio tapes, and disks will be kept in a locked file cabinet in the Department and the key to the cabinet will be kept in a separate location. In compliance with the Federalwide Assurance Code, data from this study will be kept for three years, after which all data from the study will be destroyed (45 CFR 46.117).

EMAIL dbub@nnu.edu

(208)477-122

A. BENEFITS

There will be no direct benefit to your child from participating in this study. However, the information they

provide may help educators to better understand how personalized learning is impacting instruction in your school district.

B. PAYMENTS

There are no payments for participating in this study.

C. QUESTIONS

If you have questions or concerns about participation in this study, you should first talk with the investigator. **Derek Bub** can be contacted via email at **dbub@nnu.edu** via telephone at (**208) 477-1229**. If for some reason you do not wish to do this you may contact **Dr. Bethani Studebaker**, Director of Doctoral Programs in Educational Leadership at Northwest Nazarene University, via email at **bstudebaker@nnu.edu** via telephone at **(208)467-8802** or by writing 623 S. University Blvd, Nampa, Idaho 83686.

Should you or your child feel distressed due to participation in this, you should contact your own health care provider.

D. CONSENT

You will be given a copy of this consent form to keep.

PARTICIPATION IN RESEARCH IS VOLUNTARY. Your child is free to decline to be in this study, or to withdraw from it at any point. Your decision as to whether or not to participate in this study will have no influence on their present or future status as a student in the **West Ada School** District.

Name of Student: _____

I give my consent for my child to participate in this study:

_____ _____

Signature of Parent/Guardian of Participant **Date**

I give my consent for the interview and discussion to be audio taped in this study:

_____ _____

Signature of Parent/Guardian of Participant **Date**

I give my consent for direct quotes to be used in this study:

_____ _____

Signature of Parent/Guardian of Participant **Date**

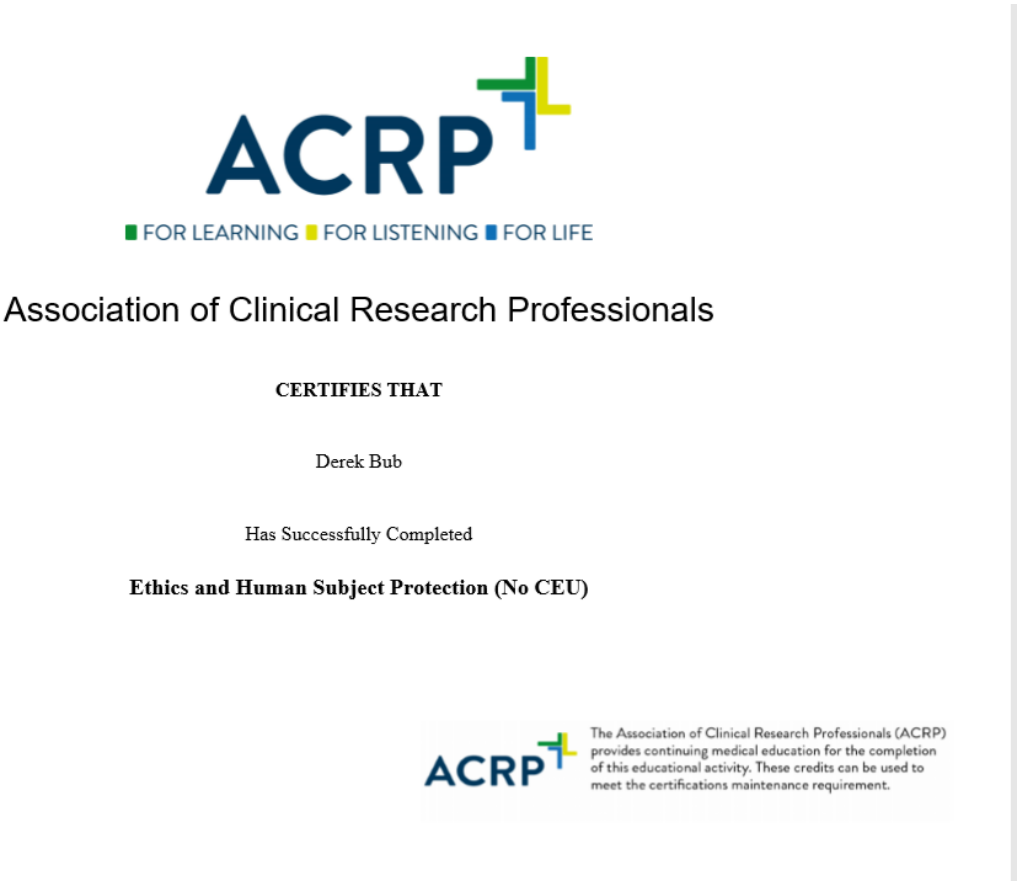
_____ _____

Signature of Person Obtaining Consent **Date**

THE NORTHWEST NAZARENE UNIVERSITY INSTITUTIONAL REVIEW BOARD HAS REVIEWED THIS PROJECT FOR THE PROTECTION OF HUMAN PARTICIPANTS IN RESEARCH.

Appendix D

Clinical Research Ethics Certification



ACRP

FOR LEARNING FOR LISTENING FOR LIFE

Association of Clinical Research Professionals

CERTIFIES THAT

Derek Bub

Has Successfully Completed

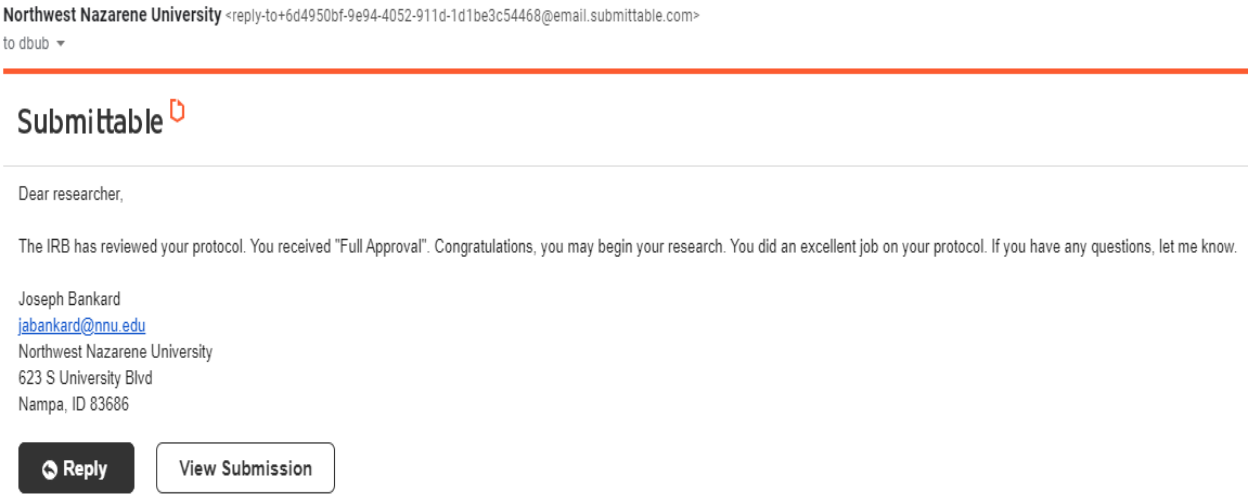
Ethics and Human Subject Protection (No CEU)

Certification Date:
January 23, 2019

ACRP The Association of Clinical Research Professionals (ACRP) provides continuing medical education for the completion of this educational activity. These credits can be used to meet the certifications maintenance requirement.

Appendix E

Institutional Review Board Approval



Northwest Nazarene University <reply-to+6d4950bf-9e94-4052-911d-1d1be3c54468@email.submittable.com>
to dbub

Submittable

Dear researcher,

The IRB has reviewed your protocol. You received "Full Approval". Congratulations, you may begin your research. You did an excellent job on your protocol. If you have any questions, let me know.

Joseph Bankard
jabankard@nnu.edu
Northwest Nazarene University
623 S University Blvd
Nampa, ID 83686

Reply | View Submission

Appendix F

Site Approval Letters

April 16, 2020

Northwest Nazarene University
Attention: HRRC Committee
Helstrom Business Center 1st Floor
623 S. University Boulevard
Nampa, ID 83686

RE: Research Proposal Site Access for Mr. Derek R. Bub

Dear HRRC Members:

This letter is to inform the HRRS that [redacted] has reviewed the proposed dissertation research plan. Mr. Bub has permission to conduct his research study with the students and staff at [redacted]. The authorization dates for this research study is for September 15, 2020 – October 15, 2020.

Respectfully,

[redacted]

April 16, 2020

Northwest Nazarene University
Attention: HRRC Committee
Helstrom Business Center 1st Floor
623 S. University Boulevard
Nampa, ID 83685

RE: Research Proposal Site Access for Mr. Derek R. Bub

Dear HRRC Members:

This letter is to inform the HRRS that [redacted] has reviewed the proposed dissertation research plan. Mr. Bub has permission to conduct his research study with the students and staff at [redacted] School. The authorization dates for this research study is for September 15, 2020 – October 15, 2020.

[redacted]

April 16, 2020

Northwest Nazarene University
Attention: HRRC Committee
Helstrom Business Center 1st Floor
623 S. University Boulevard
Nampa, ID 83686

RE: Research Proposal Site Access for Mr. Derek R. Bub

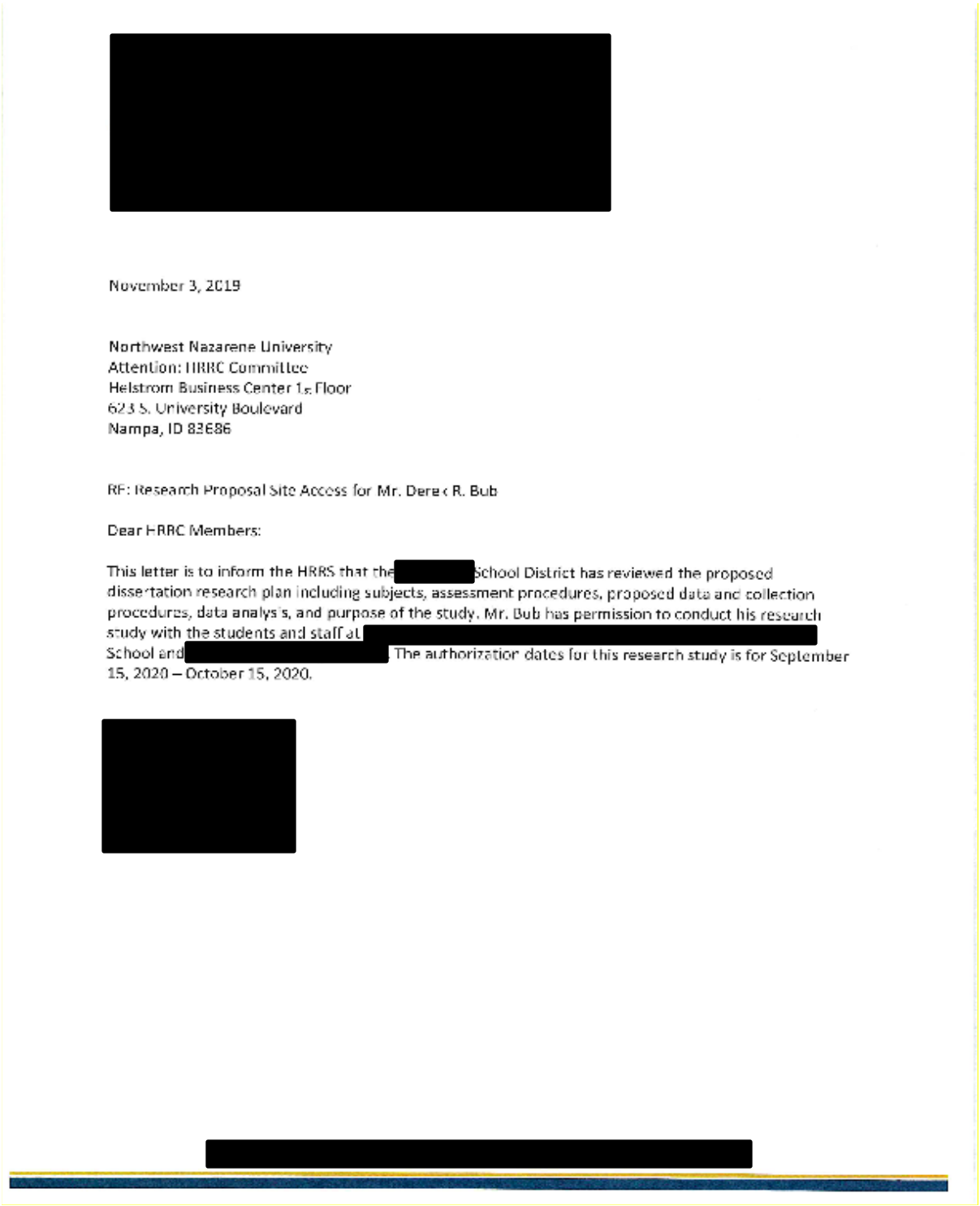
Dear HRRC Members:

This letter is to inform the HRRS that [redacted] has reviewed the proposed dissertation research plan. Mr. Bub has permission to conduct his research study with the students and staff at [redacted] The authorization dates for this research study is for September 15, 2020 – October 15, 2020.

[redacted]

Appendix G

Board Approval Letter



[redacted]

November 3, 2019

Northwest Nazarene University
Attention: IRRC Committee
Helstrom Business Center 1st Floor
623 S. University Boulevard
Nampa, ID 83686

RE: Research Proposal Site Access for Mr. Derek R. Bub

Dear IRRC Members:

This letter is to inform the IRRS that the [redacted] School District has reviewed the proposed dissertation research plan including subjects, assessment procedures, proposed data and collection procedures, data analysis, and purpose of the study. Mr. Bub has permission to conduct his research study with the students and staff at [redacted] School and [redacted]. The authorization dates for this research study is for September 15, 2020 – October 15, 2020.

[redacted]

[redacted]

Appendix H

Qualitative Interview Questions

Student Interview Protocol

- Tell me about the grading practices in this class as compared to prior English classes that you may have taken.
- Did you have any difficulty adapting to this new grading philosophy? Why or Why not?
- Has this new philosophy changed your mindset towards learning? Why or Why not?
- Compare this philosophy to the grading in other classes. Would standards-based grading work in your other classes?
- How does standards-based grading impact your mindset when it comes to accomplishing difficult tasks?

Appendix I

Email to teachers requesting to use class for study

Good Afternoon,

My name is Derek Bub, and I currently serve as the principal at Centennial High School. I am working on completing my Ph.D. through Northwest Nazarene University, with a focus on standards-referenced grading. This study specifically looks at the impact of standards-referenced grading practices on the academic self-efficacy of students. I received written permission from (Site Principal), and from the (Specific District) District Board of Education, to use classes at (School Site) for this study. Students will only need to complete a short, one-time survey that I would administer live. I am seeking grade 11 and 12 ELA teachers that are currently using standards-referenced grading practices. For the purposes of this study, these practices are defined as:

- Courses have identified priority standards and proficiency scales in use.
- Extra-credit is not given
- Students are given multiple opportunities and options to demonstrate proficiency
- Student's grades are not penalized for late work
- Final grades are based on evidence of standards mastery rather than an average

I know you are extremely busy right now; please rest assured that this will not be an extra burden on the teacher, and it will only take up about 10 minutes of class time. Would you please send me the names of a couple of teachers willing to open up their classrooms for me? I am happy to meet with them in person or over teams to address any questions they might have about the study. Thank you for your consideration, and thank you for your dedication to our students during these challenging times. If you have any questions, please feel free to contact me directly at (951)733-8213.

Appendix J

Northwest Nazarene University

Interview Protocol

Project Title: The Impact of Standards-Based Grading on the Academic Self-Efficacy of High School Students

Principal Investigator: Derek Bub, doctoral student at Northwest Nazarene University

Hi my name is Derek. If you have any questions about what I am asking you, you can stop me at any time.

I want to tell you about a research study we are doing. In this study, we are investigating the impact of standards-based grading on the self-efficacy of high school students.

You are being asked to be in this because you are a high school student participating in an English class that is graded in a standards-based grading model.

You will be asked five questions to guide our discussion. My intent is to hear from each of you on each question. I will be video recording this session so that I can have it transcribed. The video will be stored in a secure location and destroyed when it is no longer needed. Your name will not be used in this study and I will be the only person that knows your responses in this interview.

Do you have any questions before we begin?

End of verbal script.

Appendix K

Permission for Tables

Redmond, Brian Francis <bfr3@psu.edu>
To Bub Derek
Retention Policy 190 Day Inbox Retention (6 months) Expires 8/25/2021

Reply | Reply All | Forward | ...
Tue 2/16/2021 3:20 PM

Hi Derek,

You have permission to use those images for educational purposes (such as your dissertation). They may not be used for any commercial use (such as a journal publication).

I wish you success on your dissertation.

Take care,
Brian

Brian Redmond, PhD
Teaching Professor and Lead Faculty for Organizational Leadership (OLEAD)
bfr3@psu.edu
Chair Commission for Adult Learners

From: Bub Derek <Bub.Derek@westada.org>
Sent: Tuesday, February 16, 2021 5:10 PM
To: Redmond, Brian Francis <bfr3@psu.edu>
Subject: Permission

Hello Dr. Redmond,
I am writing to request permission to use The Social Cognitive Theory- Process of Goal Realization diagram for my doctoral dissertation. I am a doctoral candidate for Northwest Nazarene University, completing a Ph.D. in educational leadership. Your response to this email is adequate for permission.

The diagram was found on this website https://wikispaces.psu.edu/display/PSYCH484/Spring+2016+~+Self-Efficacy+and+Social+Cognitive+Theories

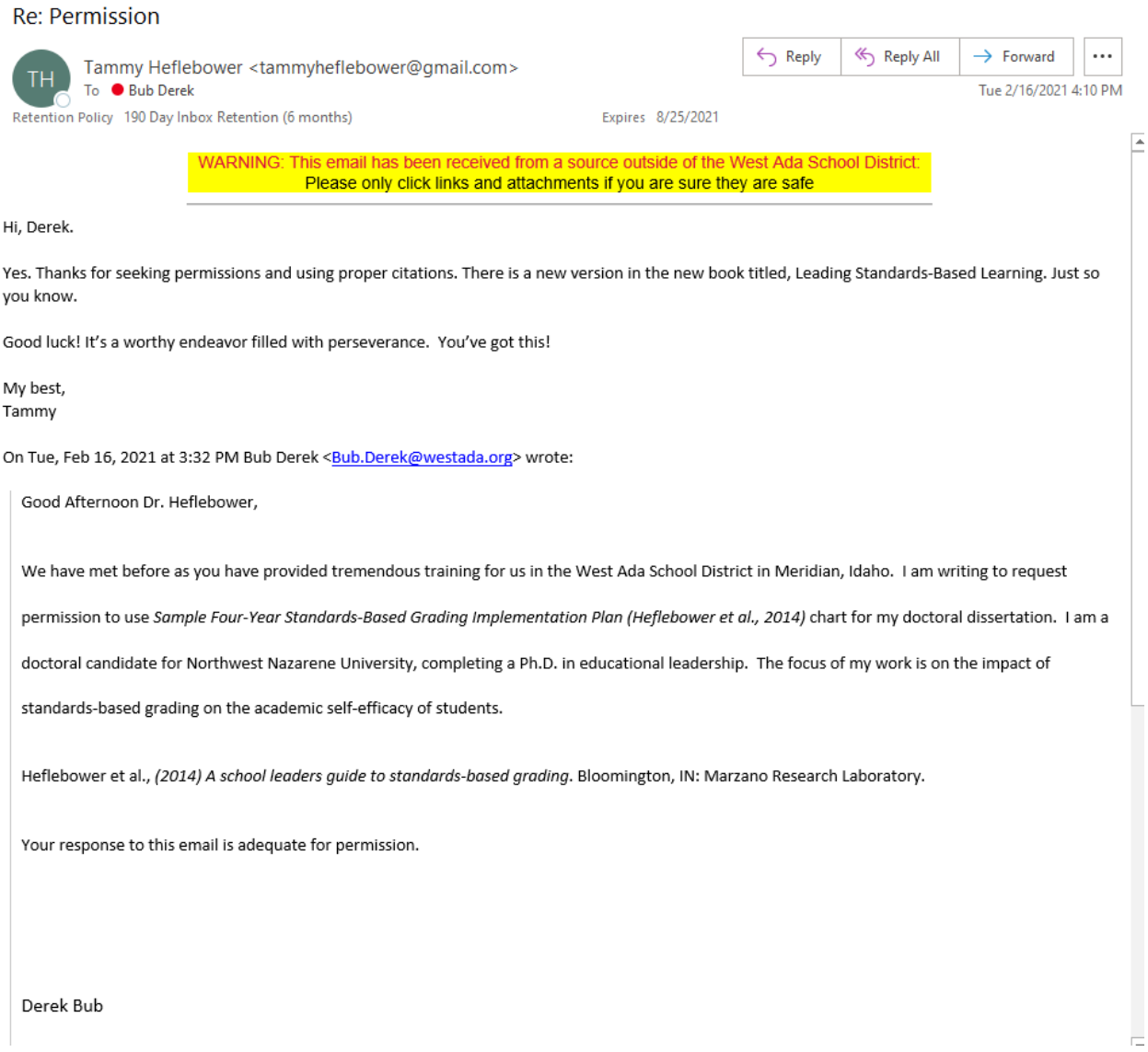
Thank you in advance for your consideration of this request.

Derek Bub
Principal
Centennial High School
bub.derek@westada.org
Office # 208-855-4250

Go Patriots!

Appendix K

Permission for Tables



Re: Permission

TH

Tammy Heflebower <tammyheflebower@gmail.com>

To Bub Derek

Reply | Reply All | Forward | ...

Tue 2/16/2021 4:10 PM

Retention Policy 190 Day Inbox Retention (6 months) Expires 8/25/2021

WARNING: This email has been received from a source outside of the West Ada School District:
Please only click links and attachments if you are sure they are safe

Hi, Derek.

Yes. Thanks for seeking permissions and using proper citations. There is a new version in the new book titled, Leading Standards-Based Learning. Just so you know.

Good luck! It's a worthy endeavor filled with perseverance. You've got this!

My best,
Tammy

On Tue, Feb 16, 2021 at 3:32 PM Bub Derek <Bub.Derek@westada.org> wrote:

> Good Afternoon Dr. Heflebower,
>
> We have met before as you have provided tremendous training for us in the West Ada School District in Meridian, Idaho. I am writing to request permission to use *Sample Four-Year Standards-Based Grading Implementation Plan (Heflebower et al., 2014)* chart for my doctoral dissertation. I am a doctoral candidate for Northwest Nazarene University, completing a Ph.D. in educational leadership. The focus of my work is on the impact of standards-based grading on the academic self-efficacy of students.
>
> Heflebower et al., *(2014) A school leaders guide to standards-based grading*. Bloomington, IN: Marzano Research Laboratory.
>
> Your response to this email is adequate for permission.
>
> Derek Bub

Appendix M

Permission for Tables

Re: Permission

Kenoc <kenoc@aol.com>
To Bub Derek

Reply | Reply All | Forward | ...

Tue 2/16/2021 6:28 PM

Retention Policy 190 Day Inbox Retention (6 months) Expires 8/25/2021

Click here to download pictures. To help protect your privacy, Outlook prevented automatic download of some pictures in this message.

WARNING: This email has been received from a source outside of the West Ada School District:
Please only click links and attachments if you are sure they are safe

Thanks for asking. As long as you cite the source you have "permission."

Stay well.

Cheers,

Ken

Ken O'Connor

Web site: www.oconnorgrading.com
Assess For Success Consulting Inc

1122-10 Guildwood Parkway, Scarborough, ON, M1E5B5, CANADA
Home - 416 267 4234; Cell (rarely used) - 416 274 1856

I would like to acknowledge that I am on the traditional territory of many nations including the Mississaugas of the Credit, the Anishnabeg, the Chippewa, the Haudenosaunee, and the Wendat peoples, and is now home to many diverse First Nations, Inuit, and Métis

-----Original Message-----
From: Bub Derek <Bub.Derek@westada.org>
To: kenoc@aol.com <kenoc@aol.com>
Sent: Tue, Feb 16, 2021 5:16 pm
Subject: Permission

Good Afternoon,

I am writing to request permission to use *Comparison of Traditional Grading and Standards-Based Grading* chart for my doctoral dissertation. I am a doctoral candidate for Northwest Nazarene University, completing a Ph.D. in educational leadership. The focus of my work is on the impact of standards-based grading on the academic self-efficacy of students.

Your response to this email is adequate for permission.

Derek Bub
Principal
Centennial High School
bub.derek@westada.org
Office # 208-855-4250