

Student Perceptions of Grading Reform at the Secondary Level

By

Scott R. Everson

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Educational Leadership and Policy Analysis)

at the

UNIVERSITY OF WISCONSIN-MADISON

2023

Date of final oral examination: 5/11/2023

This dissertation is approved by the following members of the Final Oral Committee:

Tina Salzman, Professor, Educational Leadership and Policy Analysis

Rachelle Winkle-Wagner, Professor, Educational Leadership and Policy Analysis

Peter Miller, Professor, Educational Leadership and Policy Analysis

Dawnene Hassett, Professor, Curriculum and Instruction

©Copyright by Scott Robert Everson 2023

All Rights Reserved

Dedication

To mom and dad—I know you would be proud. You are forever loved and missed.

Acknowledgements

Thank you Ellen, Drew, and Wesley for your love, support, patience, and understanding. I love you more than I can express in words.

Thank you to my dear friend Treys for being my inspiration, and providing me that occasional motivation over the past two years at times when I needed it most. God I admire you.

Thank you Laurie for being a friend, a mentor, and a thought partner. Your encouragement is the primary reason I got this dissertation adventure off the ground. Your support and understanding of everything I have gone through the past year will never be forgotten.

Thank you Brian for being a friend and an amazing editor. Your feedback, support, questions, and grammatical expertise are greatly appreciated.

Thank you to my committee members: Dr. Winkle-Wagner, Dr. Miller, and Dr. Hassett. I greatly appreciate your time, your willingness to listen to me, and your feedback as my final Oral Defense Committee.

Most importantly, thank you Dr. Tina Salzman for your many years of guidance, consultation, mentorship, and friendship. What seems like a lifetime ago, you taught me how to be a school leader and an administrative professional. More recently, you have taught me how to persevere, process feedback, sift and winnow, focus my thoughts and writing, and to make sense out of this dissertation adventure. Your support has been nothing short of incredible.

Table of Contents

Dedication	i
Acknowledgements	ii
List of Figures	viii
List of References	ix
Abstract	x
Chapter 1: Introduction	1
Statement of the Problem	3
Research Questions	4
Definition of Terms	4
Chapter 2: Literature Review	6
Contemporary Grading Practices	7
Grading Practices Aligned to Purpose	11
Grading and Feedback	11
Academic Achievement	12
Practice Work Without Penalty	12
Progress, Product, and Process	13
Grading Practices Counter-indicative to Purpose	13
Grades as Necessary for Learning	14
Grading on a Curve	15
Grading Homework	16
Grading Behavior	17
Grades as Punishment	18
Grading on a 100-point Scale	19

Grade Point Average and Class Rank	19
Mean Averaging Grades	20
Using Zeroes in Mean Average Grading	20
Grade Inflation	21
“Hodgepodge” Grading	22
Grading Reform	23
Standards-based Grading	25
Stakeholder Perceptions, Beliefs, and Biases in Grading	29
Teacher Autonomy in Grading	30
Teacher Perceptions in Grading	34
Family/Parent Perceptions of Grading	37
Student Perceptions of Grading	39
Student Voice in Education	41
Literature Review Conclusion	42
Chapter 3: Research Methodology	44
Purpose of Study	44
Conceptual Framework	45
Research Methods	47
Qualitative Research Design	47
Comparative Case Study	48
Site Selection	49
Participants	51
Data Collection Methods	52
Student Interviews	53
Consent Procedures	55
Interviews with Administrators and Teachers	55

Observations	56
Documents	57
Data Analysis	58
Trustworthiness	60
Ethical Considerations	60
Positionality	62
Pilot Interviews	62
Chapter 4: Presentation of Research and Findings	64
Case One: Northeastern High School	64
School Context	64
Northeastern High School Community	65
NHS Transition to Targets-based Grading	66
Study Participants	70
Administrator Perceptions	71
Administrator Background Experiences and Beliefs	72
Administrator Perceptions of Fidelity of Implementation	74
Administrator Perceptions of Student Experiences	75
Teacher Perceptions	76
Teacher Background Experiences and Beliefs	77
Teacher Perceptions of Student Frustration	80
Teacher Perceptions of Overall Implementation of Targets-based Grading	82
Student Perceptions	85
Student Comfort and Preferences with Grading	85
Student Perceptions of Grading Related to Post-Secondary Opportunities	87
Student Perceived Value of Targets-based Grading	88
Student Perceptions of Multiple Opportunities of Assessment	89

Student Confusion with Targets-based Grading	90
Student Perceptions of Teacher Change	92
Student Perceptions of Teacher Inconsistency	93
Case Two: Southeastern High School	94
School Context	94
Southeastern High School Community	95
SHS Transition to Targets-based Grading	96
Study Participants	99
Administrator Perceptions	100
Administrator Background Experiences and Perceptions	100
Administrator Perceptions of Implementation	100
Administrator Perceptions of Student Experiences	102
Teacher Perceptions	103
Teacher Background Experiences and Beliefs	104
Teacher Perceptions of Student Experiences	106
Teacher Perceptions of Initial Struggles of Implementation	108
Teacher Perceptions of Growing Acceptance and Shifting Culture	110
Teacher Perceptions of Legacy Practices	110
Student Perceptions	112
Student Perceptions of Positive Impact of Targets-based Grading	112
Student Perceptions on Retakes and Reassessment	114
Student Perceptions of Homework and Legacy Practices	115
Student Perceptions of Change in Teacher Practices	116
Student Preferences with Grading Systems	117
Chapter 5: Discussion and Conclusions	121
Themes of Findings	121

Comparison of School Contexts and Demographics	123
Structural/Administrative Comparison	123
Comfort	126
Familiarity, Confusion, and Frustration	127
Grading Preferences	129
Teacher Perceptions of Student Confusion and Frustration	130
Teacher Perceptions of Student Preferences	131
Academic Relevance	132
Student Perceptions of Grading Reform Impacting Their Learning	133
Student Perceptions of Reassessment Practices	135
Teacher/Administrator Perceptions of Grading Impact on Students	136
Impact of “Hodgepodge” Grading Practices on Students	137
Perceived Fidelity of Implementation	138
Student Perceptions of Implementation	139
Teacher/Administrator Perceptions of Fidelity of Implementation	139
Student Perceptions of Teacher Autonomy	140
Student Perceptions of Actual Change	141
Conclusion	142
Revisiting the Conceptual Framework	145
Limitations of the Study	147
Implications for Practice	149
Implications for Further Research	151
Significance	151

Figures

Figure 3.1: Conceptual Framework:Secondary Grading Practices Continuum of Change	45
Figure 4.1: Demographics for Northeastern High School	64
Figure 4.2: Targets-based Learning Professional Development	67
Figure 4.3: Northeastern High School Grading Proficiency Scale	68
Figure 4.4: Communication to Families	69
Figure 4.5: Northeastern School District Website Rationale for Change	70
Figure 4.6: Northeastern District Grading Policy	70
Figure 4.7: Identification of Northeastern Teacher and Administrator Participants	71
Figure 4.8: Identification of Student Participants	71
Figure 4.9: Iceberg Metaphor of Targets-based Grading and Learning	73
Figure 4.10: Staff Meeting Slide Identifying Grading Inconsistencies	75
Figure 4.11: NHS Staff Meeting Prompts Addressing Staff Pushback	84
Figure 4.12: Demographics for Southeastern High School	94
Figure 4.13: Proficiency Criteria for Southeastern High School Targets-based Grading	97
Figure 4.14: Southeastern High School Rationale and Operations of Targets	97
Figure 4.15: Identification of Southeastern Teacher and Administrator Participants	99
Figure 4.16: Identification of Student Participants	100
Figure 4.17: Southeastern High School FAQ Addressing Extra Credit	111

Figure 4.18: Reassessment for Grading Expectations	114
Figure 4.19: Tenets of Targets-based Grading	120
Figure 5.1: Primary Themes of Findings	122
Figure 5.2: Comparison of Structural/Administrative Variables	123
Figure 5.3: Comparison of Variables of Comfort	127
Figure 5.4: Comparison of Variables of Academic Relevance	133
Figure 5.5: Comparison of Variables of Perceived Fidelity of Implementation	138
Figure 5.6: Revised Conceptual Framework:Secondary Grading Practices	146
References	155
Appendix A Student Interview Questions	169
Appendix B Administrator Interview Questions	171
Appendix C Teacher Interview Questions	173
Appendix D Observation Protocol - Building Environment and Activity	175
Appendix E Observation Protocol -Teacher Team Meetings	176

Abstract

This paper presents a qualitative, multisite comparative case study from two midwestern high schools investigating high school stakeholder perceptions of grading practices experienced while transitioning from a system that primarily utilized historical grading practices to a more contemporary, reformed system implementing an iteration of Standards-based grading. The review of literature provides a synthesis of the current state of grading practices and reform efforts that intersect with historical, traditional practices as well as literature related to stakeholder perceptions. Three primary themes emerged from the findings across both sites as it related to perceptions of implementation and transition of grading practices: comfort, academic relevance, and fidelity of implementation. The findings and analyses primarily focus on a student perception perspective of school interest and the school experience as well as student perception whether the grading reform transition still held on to previous, legacy grading practices.

Chapter 1:

Introduction

“Why would anyone want to change current grading practices? The answer is quite simple: grades are so imprecise that they are almost meaningless.” - Robert Marzano

Grading has been an integral part of the US education system for as long as public education has existed. At about the turn of the 18th century, Yale University implemented a grading system using four broad markings which in turn led to a four level system. According to Durm (1993), this system was the precursor to the four point system that still exists in many secondary schools and universities to this day. In the late 1800s, Mount Holyoke College adopted a grading system that was the precursor to the A, B, C, D, F system synonymous with middle and high school education in the US today. In the 19th century, parents were oftentimes given oral reports from the teacher via home visits. Eventually, oral reports were replaced by a written narrative report on primary subjects such as arithmetic and reading, which came to be known as the report card. (Durm 1993, p. 2-3). In the 20th century, subject specific grading was implemented as a means to navigate and evaluate the world of student progress (Curreton, 1970). Since its inception, grading in public education has been a mechanism to inform students how they were doing, sort and rank students, and group students for instruction (Curreton, 1970; Durm, 1993; Vatterott, 2015).

Although grading has been a prevalent component of the education system for well over 100 years, it largely remains unchanged, and fraught with problems. Finkelstein (1913) wrote the

following with regard to the then-utilized A, B, C, D, and F grading system which even then recognized the need for grading reform:

When we consider the practically universal use in all education institutions of a system of marks, whether numbers or letters, to indicate scholastic attainment of the pupils or students in these institutions, and when we remember how very great stress is laid by teachers and pupils alike upon these marks as real measures or indicators of attainment, we can but be astonished at the blind faith that has been felt in the reliability of our marking system. School administrators have been using with confidence an absolutely uncalibrated instrument. What faults appear in the marking systems that we are now using, and how can these be avoided or minimized? (p. 1)

Fair or otherwise, contemporary grading practices are often more based on tradition than evidence of actual effectiveness (Guskey 2020).

It is important to note that grading and grading practices in K-12 education does not exist in and of itself. It is an outcome that overlaps and is juxtaposed with the complex, massive topics of curriculum, instruction, assessment and feedback. However, the focus of this study is solely on grading and the role of grading including its relationship to grading reform and stakeholder perceptions of grading.

The push for grading reform and to evolve grading practices in K-12 education to a Standards-based grading model is not cutting edge. In fact, largely as a response to the implementation of the Common Core State Standards, authors and scholars have recommended a paradigm shift to a Standards-based or Targets-based grading approach (Guskey, 2009; Hierck and Larson, 2018; Iamarino, 2014; Marzano, 2010; Moss and Brookhart, 2012; O'Connor, 2017; Westerberg, 2016). The overall literature on Standards-based grading implementation and

recommendations from practitioners and theorists surrounding the implementation is vast; however, empirical literature is almost nonexistent. Brookhart et al. (2016) indicated “studies of SBG (Standards-based grading) to date have focused mostly on the implementation of SBG reforms and the relationship of SBG to state achievement tests designed to measure the same or similar standards” (p. 828). This implies there is a need for research in Standards-based grading well beyond the myopic focus of standardized achievement tests.

Statement of the Problem

Empirical literature addressing stakeholder perceptions of grading particularly to the population of which we inflict our grading practices upon—students, and by proxy their families/parents—is quite limited. Further evidence of the lack of significant empirical studies around student and parent perceptions of grading and grading reform is supported by Townsley and Buckmiller (2020) “The literature consists of a combination of essays about the problem of classroom grading, surveys of teacher perceptions of their grading practices, and some empirical studies examining the effectiveness of various grading practices” (p. 6). Research delving into the impact that grading reform has on students is warranted and necessary. It is imperative that public education garners a better understanding of the perceptions and beliefs, especially that of students, as it pertains to the grading practices of which students’ are involuntarily immersed. Students’ perceptions of clarity and communication related to Standards-based grading and its effect on learning and the school experience is absolutely worthy of further investigation. This study is essential because it provides comparative insight to student understanding of and experiences with a previous, traditional grading system and a newer, reformed Standards-based Grading system, across multiple sites.

Research Questions

The purpose of this research study is to better understand student stakeholders' perspectives of grading reform in the transitional stages of moving from a traditional secondary level grading system into a Standards-based Grading system. The research questions that guide this study are as follows:

1. How has transitioning to Standards-based Grading reform affected students' levels of perceived interest in academic coursework and school experience?
2. From a student perspective, what "legacy practices" in grading still exist, if any, within a Standards-based Grading system?

Definition of Terms

This list includes a selection of key terms and concepts related to grading and grading practices at the secondary level.

Criterion-referenced Grading - Grading practices where a student performance is evaluated against a predetermined criteria, level, standard or scale.

Grading - The act of applying a standardized measurement, generally done by a teacher, of varying levels of achievement to a task, assignment or assessment, or in a class or course.

Grades are typically assigned via a letter, as a range, a percentage, or a number.

Norm-referenced Grading - Grading practices where student performance is evaluated in relation to the performance of other students in the class/grade or predetermined population of students.

Reform Strategies - The process of analyzing grading practices and purposes and its effects on student learning, and shifting to alternative systems.

Stakeholders - Anyone who has a vested interest in the success of a school or school system (including grading).

Standards - Learning targets, goals, objectives and/or standards are concise, written descriptions of what students are expected to know and be able to do at a specific stage and/or grade of their learning.

Standards-based (Targets-based) Grading - Grading practices that are aligned to the specific written learning target, goal, objective and/or standard and indicate progress of mastery/proficiency toward the specific standard.

Traditional Grading - Grading practices/scales that take into account factors oftentimes not related to academics and are overall subjective in nature. Traditional grading is usually based on a calculation of numbers and percentages.

Chapter 2:

Literature Review

The purpose of this literature review is to examine the literature of grading practices and purposes at the secondary level in public education, grading reform strategies, and stakeholder beliefs, biases and perceptions related to grading. These topics are relevant and current in educational research—after all, grading is a commonality across tens of thousands of classrooms, schools and educators throughout the nation, and grading practices impact hundreds of thousands of students and their parents every year if not every school day.

I initially conducted the literature review process through journal article searches primarily in the ERIC, JSTOR, and Google Scholar databases. Initial search terms (traditional and Boolean) used were “grading,” “grading reform,” “perceptions AND grading,” “beliefs AND grading,” and combinations/variations of terms. The goal of my search was to obtain a historical overview as well as contemporary findings related to K-12 (with an emphasis on secondary level) grading practices and purposes, grading reform, teacher, administrator, parents, and student beliefs and biases, and the juxtaposition and overlap of these topics. I then searched through targeted article references to identify additional articles and texts for review.

The scope of the initial search ranged from 1990–2021, although the continued citations search led me to relevant articles and texts throughout the 1900s. My initial intent was to capture more recent, contemporary empirical findings and research; secondarily, I wanted to ensure that the scope of the review reflected the historical angle of the grading topics as well. I believe the saturation point occurred when I was led back to the same studies and authors I had previously reviewed.

Contemporary Grading Purposes

In its simplest term, a grade is a mark indicating a degree of accomplishment in school. Grading is the measure of performance that bridges classroom teaching and student learning, and ideally, should be a measure of learning toward a pre-established outcome. Consensus on the purpose of grading practices in public education is far from established amongst not only educators, practitioners and scholars, but amongst the general public as well. Reeves et al. (2017) stated that grading policies are still the “wild west of school improvement, in which policy coherence is more apparent in claims than in practice and anyone armed with a red pen can make decisions with devastating instructional consequences” (p. 42). A prominent example of this “wild west” in the empirical literature was highlighted as McMillan (2001) surveyed almost 2,300 grades 6–12 core academic teachers from 69 middle and secondary schools and found that grading practices included an inconsistent, multitude of patchwork factors broken down into four general categories: academic achievement; academic enablers including effort and participation; external benchmarks; and extra-credit and fringe contributors. Likewise, other studies have indicated that contemporary secondary level grading practices have included wide ranges of purposes, practices, beliefs and strategies, even going as so far to indicate grading includes a “hodgepodge” approach as well as everything “including the kitchen sink” (Cizek et al., 1996; Cross & Frary, 1999; Duncan and Noonan, 2007; Frary et al., 1993; Liu, 2008). Further reinforcing the idea that grading serves a wide variety of purposes, Olsen and Buchanan (2019) state the following:

Nowadays, teachers assigning grades to students in classes serves many purposes. It is an objective measure of a student’s academic achievement or mastery. It is a teacher’s

subjective estimation of the quality of a student's work -- or effort, or ability, or willingness to follow directions. It is a way of comparing students with each other. It is a tool to inspire. It is a tool to control. (p. 2005)

There is no shortage of scholarly and practitioner publications regarding grading philosophy and suggested best practice. Despite this outpouring of publications and commercially-marketed products related to grading over the past 20 years, the amount of contemporary empirical literature surrounding grading is quite limited (Olsen & Buchanan, 2019). One explanation for a lack of empirical research is as follows: "the boundaries of scholarship around grading are ill-defined because grading students is linked to broader notions of assessment and even habits of mind, growth-versus-fixed mind-sets, and other perspectives on cognitive development" (p. 2034). The point the authors are making is that simply put, grading is a complex process that involves a variety of variables related to cognition and thinking and subsequently, grading *should* be a fluid process immersed in growth and learning. Instead, however, traditional grading in public education tends to be a singular, static event culminating in a score that marks the end of a finite period.

Grading has served a multitude of purposes in traditional and contemporary education (Airasian, 1994; Brookhart, 2011; Brookhart et al., 2016; Munoz & Guskey, 2015; O'Connor, 2011; Reeves, 2011). Brookhart (2011) suggests schools and school district grading policies should follow two primary principles:

1. Grades should reflect student achievement toward intended learning outcomes.
2. Grading policies should support and motivate student effort and learning. (p. 4)

As important as grades are to many constituents, there is not necessarily universal agreement or understanding regarding the purpose of those grades, even amongst scholars and authors. Munoz

and Guskey (2015) posit that, “All grading and reporting should start by having a clear purpose, followed by an in-depth understanding of the various criteria that can be used...ultimately, grading and reporting are other important tools for what matters most: improving student learning” (p. 68). Reeves (2011) suggests yet another view of grades in that for grades to have purpose, practices should adhere to four general boundaries that meet the respective criteria:

1. Accurate—The grade must reflect the performance of the student.
2. Fair—The grade must not be influenced by gender, ethnicity, socioeconomic status, political attitudes, or other factors unrelated to academic performance.
3. Specific—A grade is not only an evaluation, but feedback. Students, parents, and teachers must understand not only what the grade is, they must also have sufficiently specific information that they can collaborate to use the teacher’s feedback to improve student performance.
4. Timely—While there is, inevitably, a ‘final’ grade that appears on an official transcript, particularly in secondary school, that is but a postscript to a very long letter. Much earlier than the final grade, students should receive a steady stream of feedback, much in the way that students in music and sports receive from coaches feedback that is designated not merely to evaluate their performance but to improve it (p. 9).

Additionally, similar yet alternative purposes were stated in a prominent assessment piece written by Airasian (1994) which suggested that school grades are primarily used for administrative purposes, to give students feedback about their progress and achievement, to provide guidance to students about future course work, provide guidance to teachers for instructional planning, and to motivate students. The fact that grades have served such a

multitude of varied, even conflicting, purposes for the past 100 years coupled with the fact that prominent scholars suggest varying purposes has helped foster an educational landscape of grading perplexity.

Akin to the lack of universal agreement on grading purpose, there appears to be little accepted consensus of actual meaning related to secondary grading. Brookhart et al., (2016) conducted a meta-analysis of grading practices through a historical lens to answer the rhetorical question “what do grades mean?” and concluded that over the course of the past 100 years, grades have represented a variety of factors that teachers value, ranging anywhere from academic achievement to effort, motivation, ability, attention, and student participation. Brookhart et al., (2016) further indicated, “most teachers’ grades do not yield a pure achievement measure, but rather a multidimensional measure dependent on both what the students learn and how they behave in the classroom” (p. 30). Ken O’Connor (2011) stated that grades need to possess the following three traits:

1. Grades need to be meaningful, and instead of organizing a gradebook around an assortment of assignments, tasks, quizzes, and tests, gradebooks should be centered around meaningful learning outcomes such as goals or standards.
2. Grades need to be consistent across teachers as there should be no “hard” or “easy” teachers within a building or system and that the grading should be consistently applied to consistent achievement. Ideally, this level of consistency would apply to an entire department, building, district, and system.
3. Grades need to support learning as more than just a collection of points as a result of “doing the work” but the quality of work associated with learning is more pressing (pp. 4-5)

Further literature on grading and the use of grades identifies multiple other purposes and meanings including student accountability, instructional planning (Marzano, 2000) and diagnosis, motivating students (Docan, 2006; Cullen et al., 1975), and information about student achievement (Tomlinson, 2005). Even with the field there is little consensus around grades and at times the message is conflicting.

Grading Practices Aligned to Purpose

Although the range of functions and purposes of secondary level grading practices is vast and wide, and without consensus, a broad theme suggests a handful of general purposes which includes a big picture of communication on performance, growth, achievement (Brookhart, 1993; Guskey, 2006, Reeves, 2011). Within these large buckets of grading purposes (performance; growth; achievement) the literature is suggestive that some practices as indicated below are conducive to and aligned to these purposes.

Grading and Feedback

The literature on student feedback is vast; summarily and specifically, providing feedback itself is positive for students and that feedback itself is actually more influential to learning than actual grading (Black & William, 2010; Chappuis, 2012; Guskey & Bailey, 2001). Furthermore, John Hattie's meta-analysis research (2009) suggested feedback is significantly more impactful for student learning than grading. Contrary to popular belief, grading is not a necessary factor for student learning. However, grading is deeply ingrained in our nationwide school culture. In order for feedback to become a powerful partner to grading practices, it needs to be corrective in nature, timely, tied to a specific criterion, and should involve student self-evaluation throughout the process (Marzano, et al., 2001).

Academic Achievement

It may seem like common sense, but schools should include academic achievement as a major factor in determining grades. Students themselves have accurate understandings of their levels of achievement: in fact, as determined by Hattie (2009), student self-reported grades are of the highest effect size (1.44) of all meta-analytic factors studied. Fortunately, academic achievement appears prominent and relevant across multiple stakeholders. In a study by Randall and Engelhard (2009), researchers examined 108 high school teachers and administered projects, written assignments, and test scores. They found academic achievement the top component of grading, which they defined as a student's current criterion-referenced academic performance in relation to the particular class. But is the academic achievement as reported by traditional grading and report cards a truly valid indicator of student achievement? Ross and Kostuch (2011) reported a significant disconnect between academic achievement grades in school as compared to government-mandated assessments, with as many as 25 percent of students performing significantly better or worse on the mandated assessments than indicated on the local achievement-based report card.

Practice Work Without Grading Penalty

Fisher et al. (2011) indicated the importance of allowing students an opportunity to practice formative classroom work without the connection to grades. Akin to a coach not keeping score in practice and only 'grading' the actual game, teachers are encouraged to do the same in the classroom. Their research highlighted a high school that agreed to not grade students' work but instead grade entirely based on students' understanding of a performance task related to a predetermined competency. As a result, the overall school GPA rose from 2.89 to 3.36. As Frey

suggested, “when practice work is part of the overall grade, students don’t take risks, and teachers don’t get valuable glimpses into their understanding”(p. 5).

Product, Progress, and Process

One means of reforming traditional grading is through providing multiple grading criteria per subject or course that would allow a more descriptive profile reflective of three types of learning criteria: product, progress, and process (Guskey, 2020; Guskey & Bailey, 2009; Guskey & Link, 2019; Guskey, 2006). A product criteria would be utilized to communicate summative assessment of students’ achievement and performance. This type of criteria would typically entail final assessment or examination scores based off of projects, cumulative tests, and other culminating evidence. A process criteria represents a continuum of learning beyond just a summative product. For instance, classroom quizzes, formative assessment, and homework would all be examples of process criteria. Lastly, progress criteria would be used to indicate educational growth over a designated period of time.

Most secondary level educators can agree that grades should reflect students’ performance as it pertains to specific learning criteria. Guskey (2011) suggests that when some of a student's grades reflect academic achievement while others represent an amalgamation of several factors, it is nearly impossible to gain a clear picture of where that student's level of academic proficiency truly lies. Utilizing the multiple criteria of product, progress, and process is a means of improving the validity, reliability, and fairness within a traditional grading system (Munoz & Guskey, 2015).

Practices Counter-indicative to Purpose

Contemporary educational grading practices include a wide range of specific practices that are not conducive to student learning, nor are they representative of valid or reliable practices (Brookhart & Guskey, 2019; Cross & Frary, 1999; Gordon & Fay, 2010; Guskey, 2011;

Guskey & Brookhart, 2019; Kohn, 1999; Kohn, 2013; Malouf, 2008; O'Connor, 2011; Stiggins, Frisbie, & Griswold, 1989 Wormeli, 2018). Guskey (2011) identified five primary practices that stand in the way of contemporary grading reform:

1. Using grades to differentiate students
2. Grading students on a “normal distribution” bell curve
3. Grades are used to judge students against their peers
4. Poor grades are used to prompt students to work harder
5. One singular grade or score used to represent an entire subject or course (pp.17-20)

Stiggins, Frisbie and Griswold (1989) found that although classroom achievement was one significant factor related to teacher grading outcomes among the 15 teachers studied, a majority of those teachers still used behavioral traits such as effort and extra credit, and there was little consistency in practice between teachers. There is also evidence that suggests that bias in traditional grading exists related to teachers' perceptions of student physical characteristics and perceptions of intelligence (Malouf, 2008).

Grades as Necessary for Learning

Whether grades themselves are actually necessary for learning is a topic that scholars of grading practices and educational practitioners have wrestled with, although the review of literature found no empirical studies related to this topic. Kohn (1999) indicated that grading itself is detrimental to students and learning. He suggested that grades themselves tend to reduce students' interest in the learning, and stated:

One of the best-researched findings is that the more people are rewarded for doing something, the more they tend to lose interest in whatever they had to do to get the

reward. Thus, when students are told they'll need to know something for a test or that something they're about to do will count for a grade – they are likely to come to view that task as a chore. (p. 59)

The argument to abolish grades altogether was relevant even almost 90 years ago (Crooks, 1933; DeZouch, 1945). Gordon and Fay (2010) highlighted the incredible variation of multiple teachers grading the same work on a 100 point scale which is suggestive that not only is the practice itself inaccurate, the subjectivity and lack of interrater reliability is a significant flaw with the 100 point system, and that grading itself is a subjective practice. Kohn (2013) stated the following:

Grades tend to diminish students' interest in whatever they're learning. A “grading orientation” and a “learning orientation” have been shown to be inversely related and, as far as I can tell, every study that has ever investigated the impact on intrinsic motivation of receiving grades (or instructions that emphasize the importance of getting good grades) has found a negative effect. (p. 147)

Kohn (1999) also stated that grades tend to reduce students' preference for challenging tasks and that the greater pressure students feel to attain a top grade, the less inclined they will be to challenge themselves. Students may adapt to an environment where what matters most is good grades and they will take the path of least resistance to attain those grades. Additionally, students may simply lose interest in what they are learning and be less apt to engage in deep, risk-free thinking. Although an educational world without grades may be desirable, it is likely unrealistic and unattainable (O'Connor, 2002; Wiggins, 1996).

Grading on a Curve

Normative criteria and grading on a curve is not helpful in informing educators how students are performing as measured against a set of targets or learning goals. Instead, it utilizes

a normal distribution curve to compare how students in a small subset are performing against each other. It makes grading highly competitive between students and deters students from helping or learning from each other and fosters unnecessary competition between students. In this environment of pitting students against students where better grades are formed based on how students fared in comparison to their peers, the literature suggests more students end up penalized than rewarded and the validity of this practice is highly questionable (Crocker et al., 2003; Johnson, Johnson, & Tauer, 1979; Johnson, Skon, & Johnson, 1980; Kulick & Wright, 2008). Overall validity and meaningfulness of final grades also becomes questionable if there is not a clear separation of how students perform related to a criterion or standard versus a norm reference such as classroom peers (Friedman & Frisbie, 1995).

Kohn (1999) suggested that grading on a curve is the single most destructive, harmful grading practice largely due to the intrinsic unfairness that goes with arbitrarily limiting the number of students who can receive a top grade. Contemporary grading practices have largely moved beyond grading on a curve and it is not generally seen much anymore in public secondary school levels. (Wormeli, 2006)

Grading Homework

A primary concern with traditional assignment of homework (O'Connor, 2011) is that “when homework assigned as practice is scored and included in grades, what becomes most important to students is that it be done because it ‘counts,’ not because of any learning that might occur” (p. 110). O'Connor (2011) also suggested that homework is more of a compliance issue than an actual learning task which even contributes to cheating in the sense that it does not matter who (e.g., parent, friend, sibling, etc.) completes the tasks so long as the task itself is returned to school completed. Some researchers have indicated that beyond improving grading

practices related to homework, that homework altogether needs to be completely repurposed, non-graded, or altogether eliminated (Kohn, 2006; Schimmer, 2016). Reeves et al. (2017) indicated the following regarding the practice of grading homework:

The compulsion to grade homework is often based on the conviction that applying a score to practice, even when done in non-ideal conditions, will lead to better performance. In fact, this approach to homework leads to two types of negative outcomes -- blindly compliant students who sullenly work at skills that rarely matter, and their even more sullen peers who work at nothing, unable to even approach the task because they can't do it independently. The first group finds school excruciatingly boring; the second group finds it humiliating. Students in neither group engage in authentic learning. (p. 43)

Grading homework also raises questions of time and equity which may further invalidate the practice of grading homework. The relatively recent but widely-followed educational mantra of suggesting 10 minutes per night per grade level can be attributed to Harris Cooper of Duke University (Duke Today, 2006). However, other studies have shown that homework has little actual positive benefit to student grades (Fernandez-Alonso, Suarez-Alvarez, & Muniz, 2015; Maltese, Tai, & Fan, 2012).

Grading Behavior

The intermingling of grades and behavioral based factors like effort, motivation, and attitude is a practice that should be avoided for several reasons. First, those factors simply cannot be clearly defined and measured (Stiggins, Frisbie & Griswold, 1989). More importantly, however, those factors are not directly associated with the purpose of grades as they relate to academic achievement. As Wormeli (2006) indicates, "If the grade is distorted by weaving in a student's personal behavior, character, and work habits, it cannot be used to successfully provide

feedback, document progress, or inform our instructional decisions regarding that student—the three primary reasons why we grade” (p. 22). Brookhart (2004) indicated that teachers should assess and report non-achievement characteristics such as participation and effort, but they just should not be graded or intertwined with grading.

Similarly, O’Connor (2011) suggests that when determining grades, teachers utilize hundreds of points of data encapsulated by a single letter or score. Inaccuracies occur when achievement is mixed in with behaviors such as participation, adherence to school rules, and classroom effort. The blurring of academic achievement with behaviors was also identified in a study by Randall and Engelhard (2010) where 516 public school teachers primarily assigned grades using district-identified achievement criteria, non-academic behavioral traits such as effort and classroom behavior still played a noteworthy role in grading resulting in what the researchers identified as a need for teacher professional development. Deddeh et al. (2010) coined the term “grade fog” to describe the grading phenomenon when content proficiency becomes convoluted with factors such as attendance, neatness, practice, organization, participation, and extra credit.

Grades as Punishment

Simply put, there is no evidence in existence that supports the use of grades as punishment, and in fact, has an adverse effect on students. Letter grades used punitively propagate students’ negative feelings about the grades, students placing blame upon themselves and students are left feeling unable to improve their performance and grades (Selby & Murphy, 1992). Instead of producing greater effort, lower grades tend to cause students to disengage from the learning process (Guskey, 2009; O’Connor, 2009). Wormeli (2018) indicates a core tenet to grading should be that grades are not used as rewards or punishments in an effort to motivate

students, suggesting that intrinsic motivation for learning is far more valuable than extrinsic rewards and/or punishments.

Grading on a 100 point Scale

Utilizing a 100 point percentage scale to indicate student performance has been a largely accepted, well-documented traditional practice in secondary level education that comes with issues (Guskey 2013). Guskey and Brookhart (2019) explain that a 100 point scale suggests that the scale itself would mean there would be 101 levels of student proficiency that could be attained, and by nature of the percentages, 60 percent (a typical score of 60 and below) of grades would indicate failure utilizing this system. Likewise, Stiggins (2004) suggested that even if the 100 point scale was scaled to a more representative A, B, C, D, F to 5, 4, 3, 2, 1 scaled ratio, much subjectivity and difficulty would still exist.

Grade Point Average and Class Rank

In most high schools across the country, grades determine academic honors and class rank, and they have a direct impact on scholarship opportunities and to a lesser degree, even college admissions (Reeves, 2011). However, Hoover (2012) suggested that class rank is an unreliable and inappropriate statistic, and that only 19 percent of college admissions processes still considered it when assessing admissions criteria. Even more importantly to note, with perhaps the exception of the top ranked student, calculating class rank does not help students achieve more or obtain any greater levels of academic proficiency (Guskey, 2014). Guskey (2014) also posited that “Rank-ordering students in every graduating class has nothing to do with developing students’ talent. Rather, it is unquestionably about selecting talent” (p. 16).

Additional research has also suggested limited usefulness of class rank and grade point average. Lang (2007) studied 232 of the 500 largest public school districts in the United States

and their respective grade point average and class rank determination procedures. The study found a range of inconsistencies between weighted and unweighted grading practices. The research concluded that the vast majority of school districts utilized flawed practices in determining GPA and class rank (Lang, 2007).

Mean Averaging Grades

Compiling several scores throughout the course of a unit and then averaging those scores to come up with a single composite score to represent a student's performance on several tasks, which may or may not be related, is common practice in secondary education (O'Connor, 2009; Westberg, 2016). A primary issue with this mean averaging practice is that oftentimes equal weight is given on the front end of the learning curve during instruction as equally as it may be on the back end after students have had opportunities for practice and adequate time to learn the task. O'Connor (2011) suggested that "grades are broken when the summary they provide of student achievement is inaccurate because the procedures used to arrive at the grade are faulty...grades may mislead when they are based on simply calculating the mean (average) of a series of scores, due to the effect of outlier scores" (p. 90). An equally significant issue is that mean averaging is just one of three measures of central tendencies (mean, median, and mode) so basic measurement error is a factor as well. As such, Westberg (2016) asserted "...averaging is evil when used to determine a final score for a single measurement topic" (p. 27).

Using Zeros in Mean Average Grading

Students receiving the grade of a zero for missing or incomplete work is a common, widely-used traditional practice that has been largely criticized as a grading practice (Marzano, 2007; O'Connor, 2011; Reeves 2004). The primary issue with this practice is that statistical outliers (zeroes), when combined with other scores, provide a score that typically fails to

accurately reflect learning. Reeves (2004) asserted that a zero is such an outlier that it has no grounds statistically or mathematically to be used in the standard conversion of a 100 point scale, and that a more appropriate floor to replace a zero would be appropriate. In a longitudinal study over seven years of a large urban high school of almost 11,000 students (Carey & Carifo, 2012), researchers found the supported use of a replacement score for a zero by using a 50 minimum grade to be a low-cost, low-risk strategy that was mathematically appropriate and led to a significant increase in passed courses. Reeves et al. (2017) supported the 50-as-minimum approach but indicated the following:

...some schools have responded with the minimum 50 grading policy. The idea is that the interval between different grade levels should be equal, and therefore the interval between D and F (60 and 50) should be the same as the other intervals between higher grades. But this inevitably leads to the retort that students are “getting 50 points for doing nothing” and school administrators and policymakers often beat a hasty retreat. (p. 44)

By raising the floor from 0 to 50, schools create a more statistically appropriate starting point that allows students to recover from zeros in a traditional points-based grading system. Wormeli (2018) indicated that “refusing to raise a zero to the upper end of the F range in a 100 point scale falsifies the grade report and pushes students away from learning, undermining the student maturation we all seek” (p. 204). Reeves (2008) even goes so far to describe the practice of administering zeros in grading as “toxic”.

Grading Inflation

One cited example of grade inflation (Erickson, 2011) highlighted a correlation between higher grades and bladder control: one classroom teacher awarded bonus points for students who

did not use their allotted bathroom passes at the end of the semester. This is an obvious, albeit laughable, very real example of grade distortion and inflation. Hoover (2012) indicated that grade inflation has become a significant issue in the college admissions process citing grade point averages and class rank as contributing factors; however, Pattison et al. (2013) found very little existing evidence of actual grade inflation in secondary or post-secondary education. The conflicting findings of these two studies suggests the lack of validity and reliability of highschool grades across different high schools and districts.

“Hodgepodge” Grading

As Cross and Frary (1999) suggested, the “hodgepodge” practice of mixing attitude, performance, aptitude, effort, and other variables is not only unreliable, invalid, and subjective, but the practice in doing so is actually largely endorsed by students. Students and parents alike may very well believe that grading *should* be a hodgepodge of miscellaneous classroom indicators as a result of our public education acceptance and promotion of this cultural practice. This would suggest that perhaps better educating parents and students—as well as better understanding their overall perspectives of grading—is an important aspect of changing systemic grading practices, in addition to educating teachers.

The aforementioned traditional grading practices are still deeply ingrained today within secondary level public education. When considering the root causes of these legacy grading practices withstanding change and remaining relatively intact over time, it is important to acknowledge the role of stakeholders in education directly part of and influenced by these practices. Teachers, parents, and students are key stakeholders that enact grading practices (teachers) and are subsequently affected by grading practices (students, and to a lesser but similar degree, students’ parents).

Grading Reform

Most teachers, parents, students, and even administrators agree that schools need better grading and reporting systems (Guskey & Bailey, 2001). Unfortunately, there is little consensus amongst and between those groups what form that should take. Grading practices and grading reform has been a frequently visited topic of educational literature for over 100 years (Brookhart et al., 2016). Townsley and Buckmiller (2020) indicated in a fairly recent publication that from 2010 to 2020, an Educational Resources Information Center system search yielded over 1600 citations related to grading; however, empirical research related to grading and grading reform is limited: “The literature consists of a combination of essays about the problem of classroom grading, surveys of teacher perceptions of their grading practices, and some empirical studies examining the effectiveness of various grading practices.” (Townsley & Buckmiller, 2020, p. 6)

Systemic grading reform tends to take the form of Standards-based grading. Standards can be defined as learning goals or learning outcomes. Standards-based grading (also synonymous with Proficiency-based grading or Targets-based grading) is the grading and reporting process aligned to the specific learning standards. Despite a lack of empirical literature regarding overall grading reform, some literature suggests key teacher change strategies that need to happen for reform to occur. In most schools and districts attempting grading reform, teachers' grading practices remained largely traditional utilizing a combination of achievement and behavior (Cox, 2011). For districts and schools seeking grading reform, Cox (2011) suggests, when the majority of teachers adopted a more collegial approach with dedicated time for teacher collaboration and data analysis, teachers were more apt to adopt a shared grading approach with colleagues resulting in positive, prerequisite conditions that allows grading reform to begin to happen. Allen (2005) further suggests that teachers need to forget their longstanding

grading beliefs that they had developed as students themselves, and that specific professional development should be happening during teacher training programs as part of their training to become teachers. Likewise, the author suggested teacher value and belief systems need to be challenged, and once challenged, teachers need to then receive explicit instruction and training regarding best practices and improved grading practices. This type of explicit challenge needs to happen first in order for stakeholder attitudes and beliefs around grading to change as indicated by the following:

Peoples' attitudes, beliefs, perceptions, or dispositions rarely change in the absence of confirming evidence. Some modest change may be possible, of course, and should definitely be sought. When presented with new ideas and supporting evidence with regard to reforms in grading and reporting, for example, peoples' attitudes may move from cynical to skeptical. Although not persuaded or convinced, perhaps they will engage in discussions and consider new points of view. But commitment to a new approach, confidence that it will work, and trust that positive results will follow almost never occur up front, prior to implementation. (Guskey, 2020, p. 71)

For successful grading reform to occur, schools need to migrate from a traditional, hodgepodge approach to one focused on a specific criterion or standard with student proficiency as the goal. However, as Kalnin (2014) indicated, implementation of proficiency-based grading will invariably be full of unanticipated complexities not unlike our traditional, contemporary grading practices we see in most secondary schools today. In an examination and comparison of components of grading across both middle and high school teachers, researchers concluded, fortunately, that the same broad professional development and training practices would be appropriate for all secondary level teachers as secondary level teachers tended to utilize similar

components and measures when grading students which included a unsurprising mix of achievement, student effort, ability, participation and attendance. However, high school teachers tended to weigh classroom behavior as a greater factor in grading (Liu, 2008).

Standards-based Grading

The push to move education to a Standards-based grading model is not cutting edge, and in fact, largely as a response to the implementation of the Common Core State Standards, authors and scholars have recommended a paradigm shift to a Standards-based or Targets-based Grading approach (Guskey, 2009; Hierck and Larson, 2018; Iamarino, 2014; Marzano, 2010; Moss and Brookhart, 2012; O'Connor, 2017; Westerberg, 2016). The literature on Standards-based grading implementation and recommendations surrounding the implementation is vast; however, empirical literature is almost nonexistent. To this point, Brookhart et al. (2016) indicated “studies of SBG (standards based grading) to date have focused monthly on the implementation of SBG reforms and the relationship of SBG to state achievement tests designed to measure the same or similar standards” (p. 828). Scriffiny (2008) suggests that although many districts adopt Standards-based grading in addition to traditional, 100 point grading, districts should be replacing the traditional grading altogether.

Although teaching has become increasingly Standards-based and we now know more than ever about how kids learn, traditional grading practices persist especially in middle and high schools and despite the scholarly approach of Standards-based grading becoming more widespread, it is not yet close to being universally accepted (O'Connor, 2018). As Munoz and Guskey (2015) indicated, grading must be fair, reliable, clear in purpose, valid, and useful, and they suggested that a Standards-based approach would accomplish these purposes. Scriffiny

(2008) identified seven primary reasons that a Standards-based Grading approach should replace a traditional points-based grading approach:

1. Grades should have meaning. Instead of traditional letters A, B, C, D, or F as a compilation of points, the letter grades should reflect a meaningful level of proficiency related to the course objectives.
2. We need to challenge the status quo. Traditional practices and beliefs such as awarding points for completing homework or extra credit need promote higher grades but do not reflect actual learning.
3. We can control grading practices. Despite not having control of myriad educational factors such as class sizes, difficult parents, funding inequities, etc., educators can and should control grading practices.
4. Standards-based grading reduces meaningless paperwork. Instead of frequent marking papers for points, formative assessments become more meaningful and help judge the proficiency of the entire class as a whole.
5. It helps teachers adjust instruction. A Standards-based gradebook provides teachers with information that is relevant to instruction and student learning, and students who struggle can have multiple opportunities to show proficiency toward a goal.
6. It teaches what quality looks like. We can help students understand real life concepts of meeting performance standards which relate to their future job and career lives. Basing grades on standards (in lieu of traditional practices) we can help students understand levels of performance and quality.

7. It's a launchpad to other reforms. By focusing on clarity related to standards and proficiency, educators can improve other instructional strategies such as curriculum, assessment, and intervention

The rationale for a Standards-based grading approach rings loud and clear amongst the scholarly literature (Guskey & Bailey, 2009; O'Connor, 2018; Schimmer, 2016; Scriffiny, 2008). The purpose of Standards-based Grading aligns squarely with the aforementioned rationale. Hany (2016) indicated the the following four criterion, which serves the purpose to clearly define student goals and accurately assess whether students have met said goals, must be met to ensure the Standards-based Grading serves its purpose:

1. The purpose of grading is to report on student achievement; grades should reflect mastery of specific criterion referenced standards.
2. A grade should accurately represent student achievement, meaning the grade should not include non-achievement factors such as formative work, lateness, responsibility, and effort
3. The grade should accurately summarize achievement, meaning standards should be weighted to reflect accurate reporting of expectations
4. Standards should be clearly communicated to students, parents, and other teachers so they are aware of expectations within the class.

Although the overall purpose for traditional grading in education remains convoluted and lacks consensus, the rationale for grading reform shift to a Standards-based Grading paradigm appears solid from an educational and pedagogical perspective. But what about the parent and student stakeholder perspective? Although there is an overall dearth in the empirical literature addressing this very question, Knaack et al. (2012) investigated the perceptions of stakeholders after a one-

semester trial of Standards-based grading practices in bilingual Kindergarten/First Grade, 6th grade, and 7th grade English classrooms. Student participants indicated that they better understood why they received the grade they received based on the Standards-based approach. There was approximately a 33 percent increase in parents indicating their teachers could better explain their grades to them compared to a pre-implementation survey. Based on this limited study, it indicates that a shift to a Standards-based grading system was fairer to students, teachers were better able to identify student strengths and weaknesses, and students were able to better monitor their grades throughout the quarter (Knaack et al., 2012). They concluded that the system was overall superior and the way to go moving forward based on their findings.

Despite the clarity for students, parents, and teachers when reform efforts around Standards-based Grading are effectively implemented, multiple layers of challenges exist. One example of this is illustrated in a study by Simon et al. (2010) where proficiency-based grading convergence evidence suggested that a third or less teachers adhered to valid grading and that significant difficulties occurred when teachers needed to merge both a Standards-based approach with a traditional 100 point percentage scale in order to meet district reporting requirements. This dilemma highlights not only the benefits of grading reform and a shift towards a Standards-based approach, but also points out barriers of reform including the misalignment of grading across levels of school district bureaucracy from the Math Department to the larger school to the larger district to the Board of Education. This study's authors also acknowledged the problematic practice of using numerical percentage grading in a Standards-based system using levels of achievement indicated by rubrics.

Although school districts might decide to engage in a Standards-based grading reform and are successful in navigating the various challenges and obstacles that may present during

implementation, that does not necessarily mean the intended outcomes will follow. In one 3-year long large district implementation study, a complete grading reform process was initiated with mixed results (McMunn et al., 2003). In this multi-year study utilizing interviews, focus groups, teacher surveys, and observational evidence of approximately 240 teachers, researchers found although teachers indicated they were largely changing their practices to reflect a Standards-based approach, observational evidence was inconsistent. The researchers (McMunn et al., 2003) concluded that although teacher willingness to adapt and evolve to a Standards-based approach was largely present, professional development and training was needed to help teachers effectively implement the practices. Additionally, it was recommended that the entire model of professional development change and that special emphasis be placed on individual feedback and coaching to teachers while they implemented these changes.

Stakeholder Perceptions, Beliefs and Biases in Grading

For grading reform to happen, educators need to understand the perceptions of all stakeholders, particularly teachers', parents', and students'. Not all stakeholders agree on the purpose of grades nor do they have the same understanding (Guskey, 2002). As Cross and Frary (1999) suggested from an often-cited study of over 300 middle school teachers and 8600 middle school students, teachers demonstrated a widely mixed and inconsistent form of grade calculating taking into consideration multiple achievement, academic, and behavioral factors suggesting a divide between scholarly recommendations of grading purpose and actual practice of teachers. The researchers also found that interestingly, the students indicated they endorsed those practices as well which suggests students tend to conform to the system to which they are subjected and that stakeholders "have never seriously thought about the negative aspects of hodgepodge grading" (p. 18). In a similar study of over 2300 teachers, 85 percent of teachers

indicated that they believed grades should take into consideration extraneous factors such as effort and participation (McMillan, 2001). Mixed with extraneous factors, Baron (1999) found little agreement among stakeholders regarding essential components (participation in class; homework; attendance; overall improvement; assessments in the form of tests and quizzes; written papers; effort; recent growth) for determining a grade. Most stakeholders (teachers, parents, and students) agreed that the written papers and tests were the most important components of determining a grade. Baron (1999) also found that teachers indicated that they believed non-achievement factors such as participation, effort and attendance should be taken into high account which suggests that teacher practice in grading is likely mixed and autonomous. These study findings all point to two significant similarities: teachers have a wide variety of beliefs and perceptions of grading, and are afforded a wide range of autonomy in executing their grading practices.

Teacher Autonomy in Grading

Teacher autonomy to implement grading practices as they deem appropriate is a contributing factor to the overall evolution—and lack thereof—of grading reform. Aside from the grading experiences they received as students in their preparation programs and from informal conversations they have with colleagues, the only other opportunities teachers generally have to improve grading practices occurs through disconnected professional development and inservice training—all of which tend to be inadequate means of teacher education and preparation in the area of grading (Frery et al., 1993). Although grading systems are principally the responsibility of administration, administrators need to eventually release authority and help staff members learn how to enhance personal power and informal authority to help guide changes in practice (Lambert, 2002). Teacher autonomy in grading correlates with a general lack of knowledge and

understanding of best practices in grading across a teaching staff which is both understandable and expected. Teacher autonomy resulting in inferior grading practices collides with grading reform because teachers take their grading practices seriously, personally, and emotionally (Schimmer, 2016).

Because teachers are so attached to their grading practices, they need to be involved in developing, improving, revising and executing their grading practices. As Reeves (2008) points out, the difference between a student making the honor roll and that student failing often depends on the grading practices of one individual teacher. Similarly, Zoekler (2007) interviewed secondary level English teachers and suggested that characteristics of 'fair v not fair' and 'good v bad' figure significantly in evaluation of student work amongst English teachers in this study further implying teacher subjectivity and autonomy in grading. This study also indicated that teacher autonomy in grading was influenced by values and beliefs of individual teachers. Furthermore, the researchers concluded that individual teacher's overall grading practices and philosophies were significantly different even within the same department within the same school building. Decisions on extra credit and similar factors that could significantly impact a grade were made autonomously by individual teachers and not universal across the department. Contrarily, Brennan et al. (2001) studied 736 8th graders across six public middle schools and found that teacher grading practices tended to be fairly equitable across students despite the complexities suggesting that some aspects of grading reform is happening. However, consistent with previous findings of teacher autonomy and subjectivity, the authors noted that grades tended to be higher than actual achievement based on the fact the grades took into consideration a variety of both cognitive and noncognitive factors such as behavior, attitude, effort and initiative.

Reeves (2008) indicated that teacher autonomy is a significant factor in the chaotic world of grading. He states that “...to reduce the failure rate, schools don’t need a new curriculum, a new principal, new teachers, or new technology. They just need a better grading system” (Reeves, 2008, p. 86). The implied message is that teachers cannot single handedly change a grading system in isolation: the need for administrative, instructional leadership is evident. However, despite some districts’ best efforts and attempts to implement grading reform has been met with resistance and backlash from parents and school boards (Reeves, 2017).

If existing literature is clear about what constitutes best practices, and what constitutes counter-indicative practices, then why does secondary level grading remain such a challenging and contentious issue? Brookhart (1993) states that classroom teachers do not always follow recommendations and they are not always comfortable with the grading process. Teachers also need to be involved in the professional learning/development and training that needs to happen to educate teachers in regards to best, and less than best, practices in grading at the middle school level. Well-designed, structured, powerful professional learning can lead to changes in teacher behavior and practice and ultimately better student outcomes (Darling-Hammond et al., 2017). Involving teachers in this type of rich, deep collegial and collaborative discussions about all facets of middle school grading practices is imperative. Teachers typically receive minimal coursework training in college prior to receiving their degree. One study indicated that over half of teachers surveyed indicated they received only “a small amount” of training in the area of grading practices (Cross & Frary, 1999).

Author and Assistant High School Principal Jeffrey Erickson (2010) coined the term “the third rail” in regards to grading practices and compared the sacred, ‘untouchable’ nature of secondary grading to that of government social security: anyone who dares try to change either

practice would be committing career suicide. He also indicated that grading is a monumentally personal and private experience for both students and teachers. To this point, Duncan and Noonan (2007) randomly surveyed a sample of 513 teachers and found that teachers tended to use a combination of both criterion-referenced grading and norm-referenced grading. Their findings further reinforce the fact that teacher grading practices tend to be subjective, individual and even unpredictable. In regards to this idea of subjective, individual and autonomous practice of grading, Wormeli (2006) stated the following:

There are some aspects of teaching that we keep in cages in hopes they will never escape. Collectively, they are the ‘elephant in the room’ that everyone can see but no one mentions for fear of reprisals. Grading practices are often this elephant. We don’t share our concerns with our own grading approach or that of a colleague’s often, and we don’t spend time with each other determining the meaning of a C, an A, or discussing what constitutes a 3.5 on a rubric.” (p. 89)

Stiggins, Frisbie, and Griswold (1989) suggested three possible reasons for grading discrepancies between purpose and practice: ‘best practices’ perceived as a matter of opinion; recommended practices are not always practical; teachers lack training in best practices. Townsley (2013) suggested that even when teachers are part of research supporting grading reform, their perceptions of fairness and personal belief systems still hinder buy-in. As Cizek (1996) suggests, grades have been reliably used to communicate important information to students and parents regarding performance; however, the large issue is the unpredictable and varied use of teacher practices in doing so. In surveying over 200 teachers, Cizek (1996) found that despite the fact that almost every single district had some semblance of a grading policy, less than half of the teachers studied knew that their districts had an actual policy, and only few of the teachers

studied were able to provide any actual details about their districts' grading policies.

Additionally, teachers reported that calculations of final grades included formal academic measures, informal achievement and non-achievement, as well as “the kitchen sink” (p. 174).

Teacher autonomy in grading has stood the test of time.

Teacher Perceptions of Grading

Of the less than robust base of empirical literature on grading in recent years, teacher perceptions of grading and grading practices is perhaps the most significant one within the limited literature (Townsend & Buckmiller, 2020). In a comprehensive, historical look at teacher perceptions, Brookhart et al. (2016) found that although teachers tended to perceive grades as a variety of multiple academic and behavioral factors, they did not tend to confound grading with student traits such as personality, gender, or socioeconomic status. Anecdotally, it is worth mentioning that Brookhart et al. (2016) did not identify race as a trait that teachers considered in grading in this study. The authors concluded that teacher values tended to be significant factors in grading in that teachers' desires to be fair to students in the sense that they did not want to fail students who showed what they believed to be adequate effort, despite whether adequate learning occurred.

Additional studies have further concluded wide-ranging teacher perceptions related to grading, grading purposes, and grading practices. Brimi (2011) suggested such findings and also highlighted inter-grader subjectivity as multiple teachers grading the same paper even found huge inconsistencies. Further reinforcing this trend, McMillan and Nash (2000) interviewed over 200 mathematics and English teachers regarding their beliefs about what grading should entail. In this study's conclusion, it was suggested that teachers believed their grading practices to be a part of a larger, individual philosophy of teaching and learning and that teachers used grading as

a means to motivate students and increase effort, engagement, participation and understanding. Surveyed teachers indicated that grading was viewed as a mechanism to individualize a component of teaching and learning based on individual students.

The literature regarding teacher perceptions of overall grading appears consistent, but what of teacher perceptions related to more progressive grading practices? Hany et al. (2016) examined teachers' perceptions of Standards-based grading as well as examined the degree to which teachers properly implemented Standards-based grading in their respective classroom environments. Their findings suggested that more experienced teachers were apt to be skeptical of a Standards-based grading approach and instead, preferred a more traditional grading approach; contrarily, they found that less experienced teachers were more likely to respond positively and buy into a Standards-based grading approach. The authors suggested that in order to address the existing gap between the perceptions and degree of buy-in between more and less experienced teachers, a variety of professional development opportunities should be considered, including outside observations at schools with successful implementation of Standards-based grading, as well as additional conversation and discussion amongst staff directly related to grading practices. Hany et al. (2016) concluded that "future researchers could possibly explore a comparison of classrooms where traditional grading is taking place as well as classrooms where Standards-based grading is common practice." This void of literature comparing/contrasting teachers and classrooms between traditional grading and Standards-based grading is relevant and worthy of further research.

It is evident that even a systemic shift within a school district from traditional to Standards-based grading by itself is not enough to make a sustainable difference when it comes to changing individual teacher practice. Adrian (2012) found that despite the staff shift of

grading reform to a building-wide Standards-based program, 19 percent of the 90 teachers studied still believed that behavior factors such as behavior and effort should be included in Standards-based grading, and 53 percent of respondents indicated that daily work was a significant factor in the grading process. Additionally, over half of the teachers surveyed indicated they were still using mean averages to average out final student scores. As indicated previously in this paper, all of these aforementioned practices are deemed traditional and not consistent with the purpose of grading in a Standards-based reform process. These conclusions suggest that despite district or building reform efforts, teachers may still have the autonomy to grade utilizing practices in which they personally strongly believe. This may somewhat be explained by McMillan (2003) findings which suggested that teachers balance the demands of external factors placed on them with their individual beliefs and values when it comes to assessment, measurement and grading—even if the entire system has changed.

If teachers hang onto legacy grading practices even throughout a comprehensive grading reform change, does this then suggest that shifts in teacher beliefs need to start earlier in their careers? Perhaps not even that is enough to change a deeply ingrained belief system around grading. Bonner and Chen (2009) surveyed 222 teacher candidates who at the time were currently enrolled in a university teacher training program. The researchers determined that despite the prospective teachers' coursework and practices that were recommended related to grading, the teacher candidates still held onto personal grading beliefs despite those beliefs being contrary to indicated best practices. In particular, they found that aspiring secondary level teachers were more apt to use grades as a punishing means to reflect inappropriate student classroom behavior; additionally, secondary teachers were apt to demonstrate "success bias" and laxity in grading beliefs which is a contributing factor to grade inflation.

There is reason to believe that individual teacher practices and beliefs related to grading can in fact positively change, at least minimally. Olsen and Buchanan (2019) conducted a case study in two large urban high schools of teachers undergoing professional development related to grading reform. In this case study, researchers investigated three primary questions: 1) How do teachers' prior experiences, professional preparation, and workplace characteristics influence their grading professional development experience? 2) How do secondary teachers' beliefs and practices change as a result of a year-long professional development designed to re-think grading? 3) How do schools', teachers', and students' experiences during the year reveal deeper complexities about grading as well as illuminate paths for teacher change? Olsen and Buchanan (2019) found that although the professional development in grading experiences overall showed positive change in grading beliefs and practices with the teachers studied, they also concluded that grading itself is complex intermingled and connected with broader concepts of teaching, learning and schooling. Olsen and Buchanan (2019) stated, "The fact that grading is linked to everything accounted for much of teachers' consternation during the year. Grading is to teaching what water is to fish: it surrounds them. And, like the water for the fish, most of the teachers had not previously noticed grading" (p. 2031). Although evidence for small, positive changes in teacher beliefs and practices may be evident, there does not appear to be strong empirical evidence in the literature pointing towards an obvious, effective, systemic path forward reforming and evolving grading systems.

Family/Parent Perceptions of Grading

Parents and families represent a significant population of stakeholders when it comes to all aspects of education, including grading practices. Parents' perceptions of grading is generally developed through what they themselves experienced as children, and parents not only know

these grading mechanisms, but these practices make sense to them (Guskey, 2002). Parents also indicate they want more communication from schools as well as clear information (Million, 1999). Swan et al. (2014) surveyed 235 parent stakeholders and found overwhelming support for parent-friendly language and clarity regarding their child's grades and grading processes. They also found inconsistencies between parents preferring traditional numerical percentage-based grades and parents preferring the attempted reform Standards-based grading reports. Parents and students tend to have similar beliefs regarding the primary purposes of grades; however, those beliefs are not necessarily consistent with those in the education community of teachers and counselors (Baron, 2000). While experiencing the implementation of a Standards-based grading approach done well, parents reported they were aware of the changes to the grading practices, but they also appreciated and preferred the additional communication as it related to student progress, success, and need for improvement (McMunn et al., 2003).

Grades are extremely important to parents and students. However, student outcomes in the form of grades and report cards have traditionally been ambiguous and inconsistent. Friedman and Frisbie (1995) found that the ambiguity in final student grades and student achievement relative to the course and related criterion becomes especially unclear starting at middle school and moving into high school. Swan et al. (2014) found that parents preferred the Standards-based grading communication and believed the Standards-based approach not only conveyed more information to parents, but higher quality information related to their students' grades. Furthermore, parents indicated that grade information in the Standards-based approach was easier to understand.

Student Perceptions of Grading

There is a dearth in empirical literature in regard to student voice in education, and this goes far beyond the singular topic of grading. Gonzalez et al. (2016) extensively reviewed peer-reviewed journals spanning a 20-year period from 1990 to 2010 and found only 49 published studies. Of these studies, they were almost exclusively qualitative research methods of which primarily were interviews and case studies. Sixty percent of the studies found were related to school improvement; 20 percent were related to teaching and learning. No studies were related to the topic of grading, specifically.

Although the empirical literature base regarding secondary level student perceptions of grading is sparse, one theme stands out: both students and teacher perspectives on grading are very important (Edgar, 2014). Students preferred a grading system that assessed their performance accurately and motivated them to work harder, and researchers concluded that it is important to consider both student and faculty perceptions when contemplating grading reform (Edgar, 2014).

Like parents and families, students represent a highly diverse stake-holder group...when the (district) considers and values students' perspectives in planning grading and report reforms, not only will students better understand the reasons for change, but they will be far more likely to support the implementation. (Guskey, 2020, p. 49)

Given how little empirical evidence exists related to secondary student perceptions of grading and grading reform, and how important it is to include student perspective during grading reform, there appears to be a significant need for further research of this aspect of reform.

More recently, Thiele (2018) surveyed middle school students from four urban middle schools regarding their perceptions of grading practices and found that regardless of which

practices were being utilized (traditional vs. Standards-based), it had little effect on students' mindset (fixed vs. growth) and that students were oftentimes confused about practices pertaining to a respective grading system. For instance, many students erroneously reported that extra-credit points were allotted in the Standards-based grading system, demonstrating the fact they were unclear about the attributes of each system. Thiele (2018) concluded that it is imperative that researchers and educators continue to study the impact that grading systems have on students and student belief systems, particularly as it pertains to a fixed versus growth mindset. Similarly, Evans and Engleberg (1988) surveyed just over 300 4th through 11th grade students regarding their perceptions of grading practices, specifically the attitudes students had about being graded, the perceptions and attributes of why students get good grades, and their overall comprehension of the (traditional) grading systems being used. The researchers found that as students got older and/or were higher achieving, the students tended to perceive that grading factors were within their control to influence. Although older students tended to have a better understanding of the grading systems, overall understanding for all students was limited. Older students also became more cynical of the grading systems and found the grades to be less valuable and meaningful. It is evident that even at the post-secondary level, grades are important to students. Additionally, these studies reinforce the fact that student stakeholder perspective and perceptions of grading is valuable and remains an untapped area within the larger grading literature. As Gonzalez et al., (2016) indicated, "student voice research is a vibrant and promising field of research that has the potential to play a key role in educational reform" (p. 17).

Not only is student stakeholder perspective valuable in terms of their role in perspective future grading reform, but grades and student perceptions of grades are powerful in that students report positive outcomes and actually benefit from knowing their grades and progress within a

course; however, traditionally calculated grades may take away from the learning as students tend to focus primarily on the final outcome grade (Docan, 2008). Perceptions of grades can be leveraged as students are more apt to see value in school work and be motivated to do well when they perceive positive incentives to complete given tasks and assignments (Cullen, 1975).

While the focus of this review is centered around K-12 public education and secondary students' perceptions of grading, it is worth noting that most empirical research related to students' grading perceptions pertain to the postsecondary, college, and university level. College students' perceptions of professor grading and grading fairness has a direct relationship to the time students felt they were allotted to prepare for an upcoming examination: the more time they were allotted, the more fair they perceived the professor grading to be (Gordon & Faye, 2010). It is worth noting the perceptions of post-secondary students in relation to grading fairness given a lack of this literature topic at the secondary level.

Student Voice in Education

Mitra (2018) operationally defines student voice as the ways in which students influence or participate in educational decision-making. The study and exploration of ways in which students can participate in the educational decision-making and policy process, and the impact this may have on academic engagement, learning, participation and outcomes for students, has emerged as a topic of interest over the past couple decades. The concept of student voice as something to be identified and captured is becoming increasingly popular in education (Hall, 2016). Conner (2022) found that high schools with a greater student-reported sense of belonging and student voice were correlated with higher levels of behavioral, affective, and cognitive engagement across multiple school variables. Another recent study indicated that in high schools where students reported higher levels of acknowledgement and recognition of student voice,

students had better attendance and greater academic achievement as compared to schools with lesser student voice reporting (Kahne, 2022). The importance of student voice in future educational policy and program implementation cannot be understated, including grading reform in secondary education.

Conclusion

Contemporary grading practices in secondary public education are foundationally composed of traditional practices, routines and strategies that originated decades ago and in some cases, even a full century or more. Surprisingly, the scholarly and empirical literature is without consensus regarding the actual purpose and functionality of secondary grading despite the fact grading has been an integral cog in the mainstream educational wheel alongside curriculum, instruction and assessment. Although far from consensus, a general understanding in the scholarly literature suggests grading purpose tends to adhere to common themes of teachers utilizing grading to communicate student performance, growth and achievement. Some practices used today are conducive and aligned to these general grading purposes (e.g., practice work without penalty); however, many common practices used today are counterintuitive to general purposes of grading (e.g., grading on a curve). The need for and push for grading reform has become a significantly more popular topic in the scholarly literature largely since the advent of the Common Core State Standards in 2010. The general theme of grading reform in the literature is centered on Standards-based grading, a practice that squarely attempts to align purposeful grading to pre-established learning targets/standards and subsequently minimize the autonomy and subjectivity related to individual teacher beliefs and biases, as well as minimizing legacy practices not conducive to purposeful grading. Grading reform implementation and execution affects a handful of key stakeholders. Teachers are the primary stakeholder in the public

education arena of grading and grading reform; however, two overlooked and under-studied stakeholders—students and parents—are key players and participants in the process of grading. As indicated previously, Hany et al. (2016) suggested that “future researchers could possibly explore a comparison of classrooms where traditional grading is taking place as well as classrooms where Standards-based grading is common practice.” This void of literature analyzing, comparing, and/or contrasting classrooms between traditional grading and Standards-based grading is relevant and worthy of further research. A dearth of literature exists not only with teachers in classrooms experiencing a transition of grading reform, but also especially from the lens of stakeholders most directly affected by reform practices: the students. Given that the literature is suggestive that even schools and districts engaging in systemic grading reform have teachers unable to abandon traditional beliefs, biases and practices in administering traditional grading, perhaps focusing on the receiving end of the grading—students and their perceptions—is a key missing link in the empirical research base and potential evolutionary variable to the future of grading reform.

Chapter 3:

Research Methodology

Purpose of Study

The purpose of this study is to explore student perceptions of grading reform and its impact on their academic experience as well as perceptions of whether traditional grading practices still exist despite their school's transition to a Standards-based grading system. This study will contribute to our understanding of high school student perceptions of grading reform through the students' lived experiences transitioning from a traditional grading system to a Standards-based grading system experienced throughout their tenure in high school. The findings of this study will provide insight from multiple stakeholder perspectives into the impact grading reform has on students' experiences and perceptions. The dearth in empirical research related to student perceptions of grading is the impetus driving this study.

Conceptual Framework

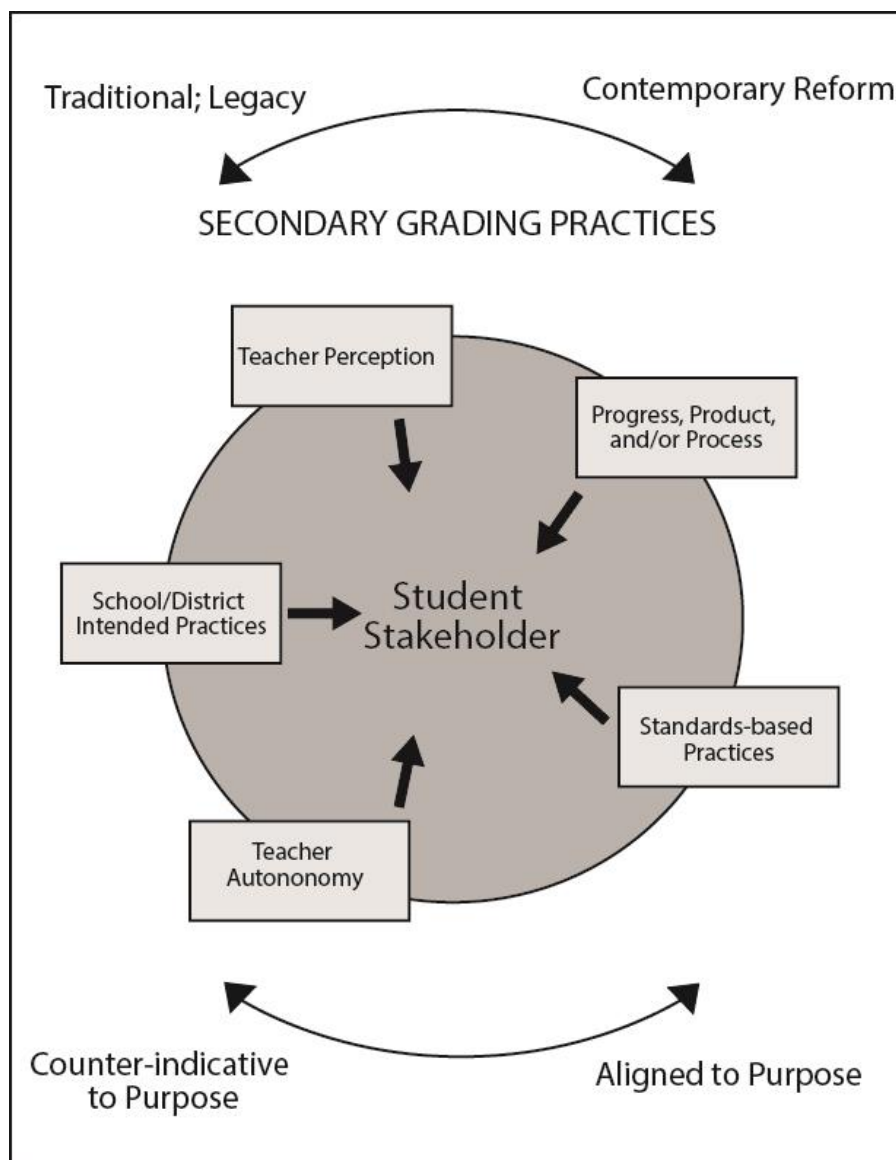
Merriam and Tisdell (2016) suggest one end of the literature review continuum is reviewing the literature to find the problem while the other end is reviewing the literature to see whether an identified problem has been previously studied. While keeping this continuum in mind and acknowledging the plethora of theoretical and scholarly literature that exists related to grading practices, I established a threshold of searching for and accepting most empirical studies related to grading that I could find due to the narrow, niche number of published studies on this topic. As I got further and deeper into the literature review, a conceptual framework started to take shape centered around the findings that student and parent stakeholder voice (student, especially) was largely absent in the empirical literature in comparison to teacher perspective.

Additionally, a related conceptual component started to materialize for me in the sense that students were at the epicenter of systemic, one-directional grading practices that have largely withstood time and change.

These findings have led me to conceptualize secondary level grading practices through the lens of a complex continuum of practice ranging from traditional, “legacy” practices that have dated back over 100 years but are still prevalent today, to a more progressive, contemporary approach that focuses on learner variability outcomes aligned to standards, targets and/or proficiencies. An additional, concurrent continuum to frame grading practices is one of overall purpose: on one end of the dichotomy, grading practices are aligned to purpose versus the opposite pole where practices are counter-indicative to purpose and problematic. Figure 3.1 below captures both continuums while positioning the student as the recipient of wide-ranging, inconsistent grading practices of which teachers and systems impose.

Figure 3.1

Conceptual Framework Secondary Grading Practices on the Continuum of Change



In this conceptual framework, the student is at the epicenter of secondary grading which represents the fact that much of what we do in education does and should center around the student and student learning. On the perimeter of the outer circle, the five boxes represent major grading practice influences to the student stakeholder: teacher perception of student performance; the progress/product/process components that makeup a grade; Standards-based reform and practices; teacher autonomy and the belief system individual teachers bring to grading; and the school or district's formal, intended grading policies and practices. The arrows

pointing toward the student stakeholder reinforce the idea of the student being the center of the five variables. The top arcing arrow represents a continuum of grading practices from (left) historical, traditional practices to (right) more modern, contemporary reform practices. Lastly, the bottom double-headed arcing arrow represents the continuum of grading practices that are counter-indicative (left) to purpose versus aligned to purpose (right) as described in the literature review. The five aforementioned boxes that represent the major grading practice influencers are positioned on the circle perimeter that is representative to their positions on the two arcing continuums as well.

Research Methods

This section details the methodology of the study including study design, participants in the study, school site specifications of the study, as well as the research instrumentality and data collection necessary to dive deeper into the confounding world of secondary level grading practices. Furthermore, this section detailed data analysis protocols, data trustworthiness, and ethical considerations.

Qualitative Research Design

As Creswell (2013) suggested “we conduct qualitative research because a problem needs to be *explored*. This exploration is needed, in turn, because of a need to study a group or population, identify variables that cannot be easily measured, or hear silenced voices” (p. 48). Particularly, the absence of student voices in the historically stagnant world of secondary level grading are the voices that need to be heard and further explored. A qualitative research design is most appropriate for the study because we utilize qualitative research “when we want to empower individuals to share their stories, hear their voices, and minimize the power

relationships that often exist between the researcher and the participants in a study (Creswell, 2013, p. 48).

Rossmann and Rallis (2012, pp.8-11) used the following table to illustrate the five general hallmarks of qualitative research as well as the five common stances of researchers who practice said research:

<i>Qualitative Research</i>
<ul style="list-style-type: none"> ● Takes place in the natural world ● Uses multiple methods that are interactive and humanistic ● Focuses on context ● Is emergent rather than tightly prefigured ● Is fundamentally interpretive
<i>Qualitative Researcher</i>
<ul style="list-style-type: none"> ● Views social phenomena holistically ● Systematically reflects on who she is in the inquiry ● Is sensitive to his personal biography and how it shapes the study ● Uses complex reasoning that is multifaceted and iterative ● Conducts systematic inquiry

The research design met all criteria for the rationale of the qualitative research as well as my role as the researcher.

Comparative Case Study

A comparative case study qualitative research approach allows the researcher to study a case (or cases) within a real-life setting in real time (Yin, 2009). Merriam and Tisdell (2016) have identified the four following characteristics as key components for understanding qualitative research which applies succinctly to case study research:

1. The focus is on process, understanding, and meaning;
2. The researcher is the primary instrument of data collection and analysis;
3. The process is inductive;
4. The product is richly descriptive (pp. 15-18)

These four tenets were the key drivers throughout the comparative case study process as I interviewed, engaged with participants individual interviews, observed stakeholders in their school environments and subsequently collected and analyzed the data, and investigated the documents (course syllabi, faculty and student handbooks, Board policies, etc.) related to grading practices within each school, department and relevant classroom. As Flyvbjerg (2011) indicated, case studies represent intensity and depth. My study entailed a deep dive attempt to get to the heart of the student perspective of grading reform. Utilizing two different yet similar high schools from two different school districts provided richer validity to the study. Creswell (2013) posited that “a hallmark of a good qualitative case study is that it represents an in-depth understanding of the case” (p. 98). The comparative study allowed for a deep understanding from multiple stakeholders at multiple sites. Merriam and Tisdell (2016) suggest that “the inclusion of multiple cases is, in fact, a common strategy for enhancing the external validity or generalizability of your findings” (p. 40).

Site Selection

The research took place within two small/medium midwestern public high schools. The following criteria were used when identifying and selecting sites:

1. Schools have high school (grades 9-12) enrollment between 200 and 999 students.
Students will currently be in 12th grade.
2. School administration allowed me in-person access to interview student volunteers (with secured student and/or guardian informed consent).
3. School administration allowed me in-person access to interview teacher volunteers with secured, appropriate informed consent.

4. School administration provided me access to all forms and documents related to policy, parent communication of grading reform, classroom and grading expectations, and any other pertinent materials including non-identifiable student documentation related to grading and grades.
5. School administration provided me access, if possible, to relevant, authentic observable staff events including team/building meetings related to dialogue, discussion, and/or decisions pertaining to grading and implementation of grading practices.
6. Schools are singular high schools within K-12 public school districts of communities no greater than 10,000 people.
7. High schools are in years 1, 2, or 3 of an implementation process that entails a formal, acknowledged shift from a more traditional, legacy grading approach to one that focuses on grading reform through a Standards-based (or Targets-based, Proficiency-based, Mastery-based, etc.) model.

I initiated my search for high schools who met my criteria by communicating via a professional network and asking whether any administrative colleagues might know of schools that fit said criteria. I received a response that provided me with seven high schools that I thought might meet the criteria upon initial information. After further review and follow-up conversations, I ruled out five schools for a variety of reasons. Three of the school sites were already at the point that they identified themselves as several years into full implementation. As such, their current high school seniors had only experienced a Standards-based Grading approach and would not meet my criteria for students who had experienced the transition from a traditional grading approach to a reformed grading approach. The other two sites indicated that they did not believe they were deep enough into the implementation process, and although they were

confident in their transitional process and eventual outcomes, they were still largely using traditional, letter-based grading practices throughout the buildings.

I narrowed the search to two specific high schools that met all criteria and eventually received approval to visit both sites on multiple occasions and interview subjects. For confidentiality purposes, the names of these two schools and all interviewees within the schools are referred to using pseudonyms. For the purpose of this study, the first school will be called “Northeastern High School” (NHS) and the second school will be called “Southeastern High School” (SHS).

Participants

A primary driver in selecting participants for this study was to ensure that all the participants have a story to tell about their lived experiences (Creswell, 2013). The primary participants in this study were students who are in their senior year and have experienced the transition of traditional grading practices to a more contemporary Standards-based Grading reform approach. I believe we have an obligation to listen to the voices of our youth. To this point Creswell suggested, “This is especially true in education, where all too often, those most affected by educational policy and programmatic decisions – the students – are absent from inquiry” (p. 161).

Participants were chosen via purposeful sampling; specifically, high school seniors, as this age group ensures the maximum amount of time for the students to have experienced high school grading practices on both sides of the continuum (traditional, legacy practices v. reform practices). As Merriam and Tisdell (2016) suggest, “purposeful sampling is based on the assumption that the investigator wants to discover, understand, and gain insight and therefore must select a sample from which the most can be learned” (p. 96). High school seniors,

specifically those who have expressed an interest in attending a post-secondary institution post-graduation, were the targeted student demographic. No students were excluded on the basis of color, ethnicity, gender, sexuality, religion, disability, and/or any other protected class. I ensured that the representative sample included a wide variety of learners, but those student volunteers were likely to be students who have experienced success in school given their intentions for attending post-secondary schooling the following year. When speaking to the building administrators ahead of time, I made them aware of potential biases related to the selection pool and ensured a diverse range of demographics (among the self-identified students intending to attend a post-secondary institution) that is representative of the overall school population.

This study also included the perspective of school administrators who were involved in leading the implementation of grading reform. The scope of administrative participation was narrow given the nature of the administrative decision making that ultimately led the site to choose a grading reform process. Nonetheless, I believed the administrative perspective would be important for this study, particularly their view of grading impact and outcomes on the students themselves. The high school administrators participated in the form of individual face to face semi-structured interviews (see Appendix B). Lastly, but equally importantly, I targeted participants in the form of classroom teachers via purposeful sampling who had firsthand classroom experience implementing the instructionally pertinent grading practices utilized in part of the larger department and/or building systems. Individual teachers play a significant role in the practical and iterative implementation of district-intended practices, including grading practices.

Data Collection Methods

Data collection in this study occurred via semi-structured individual interviews for students, administrator interviews, and teacher interviews. Merriam and Tisdell (2016) define a semi-structured approach as one somewhere between a structured and unstructured approach where the questions are flexibly worded with a desired outcome but the order and exact wording are not precisely predetermined.

Student Interviews

The interview is a powerful tool of qualitative research and in some studies, all of the data is collected via interview (Merriam & Tisdell, 2016). The individual, one-on-one student interview format has the potential to be a relaxed, comfortable and positive atmosphere where participants enjoy sharing their perceptions, ideas and experiences (Krueger & Casey, 2008). I targeted six high school students per site (although one student was unexpectedly absent the day of the interviews at one site).

Prior to my fieldwork, I requested that administration assist me in identifying willing student participants for student interviews using a natural, proportional distribution across the 12th grade student body in terms of student ethnicity, disability status, gender, etc. I focused on students who had self-identified as intending to attend college/post-secondary educational institutions the following school year. The rationale for this demographic of students planning on attending post-secondary education is that students who place a relatively high value in their academics will likely have a higher vested interest in grading practices and grading outcomes. As such, the distribution of students in regards to student academic achievement and success was determined by cumulative grade point average minimum of 3.0 on the traditional 4.0 scale. Despite the inherently arbitrary nature of grades and the shortcomings of grade point averages, it was imperative that the students interviewed represented a range of students' success that is

proportional to the distribution of the student body intending to attend a post-secondary institution. I believed that not only did this provide more valid data but it accounted for a degree of selection bias when drawing from the pool of students. The timeframe for individual student interviews was 20 to 40 minute range with 40 minutes being a firm cut off. No interviews surpassed 40 minutes. Most interviews took 15 to 25 minutes each.

I adhered to a semi-structured interview structure and the overall format of my interview questions fit within the proverbial funnel and subscribed to easing into a more indirect approach. Individual student interview questions started scripted with broad inquiry, then became narrowed in and focused on the more precise components of the grading reform objective.

Prior to all interviews across both sites, I ensured the room environment where the student interviews took place was a typical and relaxed environment for students. Both sites satisfied this criteria. I started with rapport-building small talk with each student to ensure they were comfortable with me, the setting, and the conditions of the interview. I utilized an interview protocol (Appendix A) that was given special attention to ensure it was presented in student-friendly language with minimal educator jargon. Krueger and Casey (2008) propose a three-stage developmental process of questioning that is especially relevant to program implementation such as the grading reform in the high schools of this proposed study design. Questions that seek participant understanding (e.g., Tell me about your experiences with...?) and evaluation (e.g., Think back to your first experiences with the new grading system. What were your frustrations?) were appropriate. Given that the questions are a core inquiry mechanism for this study, I was mindful to craft them in a way that represented the central research questions. I gave special attention to the questions so that they were broad enough to allow students to think and reflect upon overall grading practices, and not elicit any potential emotion around any individual

circumstances they may have experienced in a particular class or with a particular teacher. Further refinement of the questions was done via reflection following the pilot interviewing process.

Consent Procedures

Prior to arriving at the sites, I arranged with the school administration for students to receive the consent forms and have adequate time to read the forms as well as communication of the purpose of this study to students' families/parents as well. Likewise, students received adequate time to reflect on my explanation of the purpose of the study and my intended usage of the data they would be providing. I obtained consent from school administration to be on site and to receive all pertinent district grading documents. I obtained consent from parents and students alike for participation in the interview processes.

I successfully coordinated with both school administration to ensure that students had an opportunity to interview in a natural, oft visited, comfortable environment within the school buildings. Students were informed that they had the right to opt out at any time, no questions asked. No students took me up on this offer.

Interviews with Administrators and Teachers

In addition to gathering data via individual student one-on-one interviews, I interviewed individual administrators who were responsible for implementation of the grading reform in their respective school buildings/districts via a semi-structured interview approach. The interviews followed a semi-structured protocol (Appendix B).

I set up the in-depth interviews at a time and place of convenience for the two participating administrators. The interviews took approximately 40 minutes each and were done individually, face to face, within the school environment of their choosing. The semi-structured

process included scripted yet open-ended questions that allowed for further investigation and exploration of responses. All interviews were digitally recorded for transcription and coding purposes. Observations and field notes were documented throughout the interview process. I found it especially helpful to take field notes in a journal immediately after each interview.

The semi-structured, in-depth teacher interviews (Appendix C) took place in the school environment of each teacher's choosing (all four chose an office conference room). Teacher selection via purposeful sampling consisted of teachers with at least three years of experience who had district longevity that ensured they taught and graded at least two years in the previous system prior to the transition to the current reformed grading system. Teachers were core area instructors who had taught required courses (e.g., English Language Arts; Mathematics). The interviews were intended to be 30 to 40 minutes in length; almost all of them ended up around 40 minutes. The interviews followed the semi-structured format of starting with a scripted template of questions but allowed for deeper inquiry into teacher grading practices. I interviewed two teachers from each site.

Observations

The data gleaned from observations is an extremely important and fundamental means of qualitative inquiry (Marshall & Rossman, 2016). As indicated previously, a tenet of a qualitative researcher is one who views social phenomena holistically. Observation is a powerful holistic tool as "you may watch physical settings, participants, activities, interactions, conversations, and your own behavior during the observation. Use your senses, including sight, hearing, touch, smell, and taste" (Creswell, 2013, p. 166). Observations of the sites themselves, as well as of the students, teachers, and administrators in their school environment, is paramount. I used an

observational protocol for taking both description, real-time notes as well as reflective notes (Appendix D).

It was important that the opportunities for observation were authentic and grounded in grading practices that ultimately affect students. As such, I sought to gain observational access to any relevant meetings around teaching, assessment, and grading practices that included grade level meetings, department-level meetings, building meetings (e.g., mandatory teaching staff meetings; building grading committee meetings), district meetings (e.g., MS and HS vertical team meetings) and any other conversations/meetings that were relevant to the conversation and implementation of grading practices. Unfortunately, however, each site yielded only minimal authentic opportunities for observations.

Documents

Obtaining relevant school documents related to grading from each site was also key to this comparative case study data collection. As should be the case with grading documentation and artifacts for my proposed study, oftentimes documents are ready-made and easily accessible (Merriam & Tisdell, 2016). The use of documents as a supplement to interviews and observations can provide rich, valuable information in qualitative research (Marshall & Rossman, 2016). I obtained and reviewed the following documents:

- Building-level and classroom grading policies in the form of handbooks, policy documents, and syllabi;
- Communication from school personnel to families and/or students regarding implementation of form grading system changes;

- Non-identifiable student grading documentation comparing grades from the previous traditional system to the more current system; no actual educational records were utilized;
- Historical grading documentation and meeting minutes/notes;
- Staff meeting communication and slides related to implementation;
- District policies related to grading;

I was able to acquire materials and documents from both districts' websites; additionally, I successfully requested and received a variety of relevant documentation from district staff that was not publicly accessible.

Data Analysis

Merriam and Tisdell (2016) suggest a simultaneous approach to data collection and analysis. I followed a Data-Analysis Spiral put forth by Creswell (2013) that is composed of the five following components:

1. Organize the data—this was the digital means by which I organized transcribed text data gathered throughout the interview and observation process.
2. Reading and memoing—after organizing data, reading transcripts and writing notes/memos in the margins was a key mechanism for a deeper exploration of the data.
3. Describing, classifying, and interpreting data into codes and themes—forming codes of data in this phase may perhaps represent the most critical component of my qualitative data analysis. The coding process entailed me winnowing data and breaking data into smaller chunks and categories. I used an inductive coding process. After initial coding, I followed up with a thorough line-by-line coding process as a second “draft” of the coding process. I used the Nvivo software to assist with the coding process.

4. Interpreting the data—during this phase of the spiral, I worked to make sense of the data by delving into the codes, finding themes within the codes, and then organizing themes in a meaningful manner.
5. Representing and visualizing the data—the last phase resulted in a visual representation of the data in the form of a table. (p. 182-188)

The student, teacher, and administrator interviews were digitally recorded and fully transcribed via Rev app. Related to step 3 of the spiral (describing, classifying, and interpreting data into codes and themes), I organized my data via a system of coding following the Merriam and Tisdell (2016) definition which states “coding is nothing more than assigning some sort of short-hand designation to various aspects of your data so that you can easily retrieve specific pieces of the data” (p. 199). I used the online software Nvivo to help with the data analysis process. Saldana (2016) suggests that coding requires a researcher’s analytic lens but also requires a researcher to be mindful of the filter that covers the lens when perceiving and interpreting what is happening in the data.

This study utilized multiple stakeholder interviews, observations, and document review data for comparison and analysis as a means of triangulation in an attempt to ensure the most credible findings possible. As Merriam and Tisdell (2016) suggest:

With regard to the use of multiple methods of data collection, for example, what someone tells you in an interview can be checked against what you observe on site or what you read about in documents relevant to the phenomenon of interest. You have thus employed triangulation by using three methods of data collection – interviews, observations, and documents (p. 245).

Individual interviews with students, individual interviews with teachers and administrators, collection of documents related to grading practices and policies, and observations provides necessary sources for optimizing the validity and credibility of this study.

Trustworthiness

It is of utmost importance that any qualitative research holds the highest of expectations of trustworthiness. Marshall and Rossman (2016) indicate the traditional components of trustworthiness include reliability, validity, objectivity and generalizability. Merriam and Tisdell (2016) indicate that “to have any effect on either the practice or the theory of a field, research studies must be rigorously conducted; they need to present insights and conclusions that ring true to readers, practitioners, and other researchers” (p. 238). The research approach of this study held true to these tenets.

One strategy of trustworthiness I employed was asking interview participants to review the transcripts to ensure accuracy and validity which subsequently helped ensure validity with the coding process. Furthermore, I engaged some participants in a member check process that ensured not only were the transcripts literally valid to their recollection, but that the intent and meaning of their input was captured accurately and appropriately. Additionally, I utilized a running record of record keeping and have preserved all my data to ensure it would be available for further future analysis, if necessary. Marshall and Rossman (2016) suggest researchers strongly consider the aspects of data triangulation, dependability, and confirmability that lend toward a more trustworthy research design.

Ethical Considerations

Creswell (2013) states that ethical considerations should be considered in every aspect of the research process: from prior to conducting the study to being on-site to collecting, analyzing

and reporting data to publishing the study. One particularly sensitive and noteworthy area of ethical consideration in the study was in regards to the fact I interviewed students. Specifically, the two components I considered were::

- comfort of students interviewing in a one-on-one situation versus a group format
- power dynamic between me and my role as an unfamiliar adult researcher with that of the students' role as participant

I have protected the privacy and identity of all student participants by using pseudonyms for students (e.g., NE1 to represent the first student I interviewed from the Northeastern site) and schools, and removing every component of identifiable student information. Along with acknowledging and giving credence to the ethical considerations, I believe soliciting the voice of students is at the heart of my research proposal. As Marshall and Rossman (2013) suggest, “increasingly, there are calls for including children’s and youth’s perspectives as relevant and insightful in learning more about aspects of their worlds” (p. 161). I used audio recordings only as that provided better student anonymity than video recording.

The topic of grading can be a very personal one for students and teachers alike. It was imperative that the student interviews allowed students to speak honestly, freely, and openly. Ensuring that the dialogue about their experiences with their academics and achievement without feeling vulnerable, and without eliciting any unnecessary emotion from past experiences, was paramount throughout the individual interviews. As such, the priority of the questions centered around individual experiences with the grading system overall, per se, and not individual experiences in specific classes and/or specific instructors.

Similarly to my approach with students, I was mindful of ethical considerations when interviewing the adult teacher and administrator volunteers as well. I ensured their anonymity

and that their identities were protected. I was cognizant of the power dynamics between the administrator and teachers within each school by ensuring all comments and contributions from teachers are not individually identifiable.

Positionality

I would be remiss if I did not recognize the personal bias I bring as a researcher. I work as a Director of Curriculum & Instruction in a small school district with a high school similar in size to those I visited and studied via this proposal criteria. I have developed an interest in grading practices throughout my career based on my personal experiences as an educator and administrator paralleled with my reflections on practices I experienced when I was younger as a K-12 student myself. My positionality as a Director (and former High School Principal) lends to a bias that might include my perceptions of what I believe is and was the best route to implementation and how students could or should have been impacted. I attempted to guard against this bias by my mindfulness associated with the field notes and recording of responses taken, as well as mindfulness while I coded the gathered data.

Pilot Interview Protocol

I conducted a pilot interview study to test my research protocols and most specifically, the individual student interview questions. An abundance of literature supports and advocates for the use of pilot studies prior to engaging in the formal research (Creswell, 2013; Marshall & Rossman, 2016; Yin, 2018). I utilized my personal connections to pilot these protocols with a couple of students; geographical proximity was a variable for selecting the two pilot participants. In addition to testing protocols, I utilized this pilot study to subsequently fine tune and revise lines of inquiry and specific questioning.

I also piloted the administrator questions with two volunteer high school administrator participants. Similarly to the student pilot protocols, I utilized my collegial network and connected with two individual administrators who, although they did not participate in the actual study, have themselves experienced some form of grading system change so as to ensure the questions are pertinent and relevant for the pilot. I made minor adjustments to the interview questions and protocol based on the pilot.

Lastly, I pilot studied the individual teacher interview questions with two high school teacher volunteers who are also part of my collegial network, and who have an existing knowledge of Standards-based grading practices. I ensured the teacher volunteers were veteran teachers. Similarly to the administrator pilot, only minor changes to the protocols were necessary upon completion.

Chapter 4:

Presentation of Research and Findings

In this chapter, I present results and narrative descriptions of the data collected via the fieldwork and research of Northeastern High School and Southeastern High School. Both schools are small (student population between 200 and 999), midwestern high schools. They are both in the recent —within the last three years—transition from a more traditional, 100 point scale, A through F letter grade grading system, to a Standards-based Grading system.

Case One: Northeastern High School

School Context

Visits to Northeastern High School (NHS) started in November 2022. NHS is part of a large, interconnected district campus that houses all levels of students K-12. Parts of the entire facility, including several middle and high school wings, were recently constructed and some parts, including elementary classrooms, were recently renovated and surrounded by new blacktop surfacing. The campus is quaint and beautiful, located on the outskirts of the rural community of 2,700 people. NHS is a rural midwestern school of 250 students of which approximately 85% of students are White and 15% of students are Black, Indigenous, and people of color. Approximately 35% of the students are economically disadvantaged. NHS demographics have largely remained consistent the past five years with the exception that they appear to be experiencing a slightly increasing enrollment trend (233 to 254) and slightly increasing students in poverty trend (29.2% to 34.6%).

Figure 4.1 *Demographics for Northeastern High School*

Figure 4.1

Northeastern High School	
Enrollment	254
Percent Open Enrollment	17.7%
Students Economically Disadvantaged	34.6%
Students who are English Learners	2.8%
Students with Disabilities	17.7%
Students who are White	87.4%
Students who are Black, Indigenous, and People of Color	12.6%

According to the state level education department's school report summary, NHS self-indicates that their students have access to a wide array of courses aligned to individual needs with the focus on preparing students for individual post-secondary planning, with classes and instructional options ranging from Advanced Placement to credit-recovery programming.

Northeastern High School Community

Upon arrival and being buzzed in by office staff and entering the building's secured entrance the first time, my first face-to-face contact was with the office administrative assistant who was initially excited to see me, then disappointed to hear that I was not a substitute teacher. I was soon greeted by the District Administrator, Sheila Larson. Meeting Larson was followed by warm introductions to all the staff members in the office as well as a tour of the district campus and buildings.

NHS looked and felt like a miniature community in and of itself. The hallways and classrooms are interconnected through a simple 'figure 8' design, giving the feeling that the office, commons, library, and classrooms were all easily accessible and centrally located. Hallways were well-lit and full of posters, current event postings, athletics achievements, and

artifacts of student work. An award-winning welding structure, the product of a student Career and Technical Education club, was proudly on display in one wing of the school building. During passing times before and between classes, conversation amongst the student body was cordial and friendly evidenced by loud voices, laughter, and pockets of individual student conversations as they moved throughout the building. Staff members were present throughout the hallways during passing times and even moved throughout the building occasionally during class times. Several staff members stood outside their classroom doors greeting students as they entered their classrooms. NHS appeared to have a positive, warm student and adult environment.

NHS Transition to Targets-based Grading

Starting in August of 2019, Larson, the high school principal at the time, initiated a series of formal professional development work and learning with staff as well as a series of staff meetings in preparation for the Targets-based Grading transition process. Larson defined their Targets-based Grading model based on the work of Hierck and Larson (2018) as a system of grading that requires teachers to grade and report against specific learning targets, and not standards as a whole. This grading model allows parents and students to clearly understand what children are expected to learn in class, and also provides them feedback on progress toward specific learning targets. Targets-based Grading does not support traditional grading practices counter-indicative to learning such as grading on a curve, grading homework, grading behavior, grading on a 100 point scale, or grading traits of student behavior. Targets-based Grading is built on the premise that students are graded based on evidence of work primarily through summative assessment as opposed to formative work and assessments, and learning tasks such as practice work and homework.

The transitional process consisted of monthly and sometimes twice monthly meetings of teachers in a building-wide whole group as well as departmental, smaller group learning sessions. The professional learning dates and tasks are seen below in figure 4.2:

Figure 4.2 <i>Targets-based Learning Professional Development</i>	
<u>Timeline</u>	<u>Activity/Task</u>
August, 2019	Departments and teams reviewed Targets-based Curriculum documents
January, 2020	Departments and teams conducted peer review/audit of learning targets; new learning targets were written; impact on report card discussion
February, 2020	Departments and teams reviewed Targets-based Grading resources; created assessment/success criteria for learning targets; individual work time; continued report card discussion
April, 2020	Staff discussion regarding pros and cons of traditional versus Targets-based Grading
May, 2020	Virtual meeting; continued work on rubrics
August, 2020	Virtual meeting; continued work on rubrics
September, 2020	Discussions of implementation of Targets-based Grading during COVID-19
October, 2020	Targets-based Grading scenario practice; reviewed research supporting Targets-based Grading
November, 2020	Virtual meeting; Focus on student opportunities for reassessment; breakout discussion of what is working well and not working well with Targets-based Grading
December, 2020	Handbook changes needed to better reflect Targets-based Grading changes
May, 2020	Reiteration with staff regarding purposes of

	grading
January, 2021	Acknowledgement that staff all in different places with Targets-based Grading; continued work with revising/finalizing learning targets
March, 2021	Staffed watched virtual sessions of Targets-based Grading conference
August, 2021	Continued discussion of what is working well and what needs to be tweaked/revised with Targets-based Grading
September, 2021	Opportunity for staff to share out anything regarding Targets-based Grading
October, 2021	Discussion on commitments and beliefs related to Targets-based Grading
November, 2021	Discussion of hypothetical grading scenarios related to Targets-based Grading

After months of planning and staff preparation, NHS’s transition to Targets-based Grading largely occurred during the 2020–2021 school year (which also happened to be the primary COVID-19 year). They refer to their shift in grading practices as Grading for Learning, which utilizes a four indicator system that ranges from Needs Support, to Developing, Proficient, and Advanced (Figure 4.3). Their Targets-based Grading for Learning system follows a general template and approach set forth in the *Grading for Impact* (Hierck & Larson, 2018) resource, which NHS administration used to establish a formal grading proficiency scale.

Figure 4.3 *Northeastern High School Grading Proficiency Scale*

From High School Student Handbook:

GRADING/REPORT CARDS

Parents have the opportunity to view their student's grades through PowerSchool, the district student information system, via the parental access web portal. Parents may request a paper copy to be picked up at the high school office, sent home with the child, or emailed.

While there is an overall traditional letter grade (e.g. A, B, C, D or F) for the course at the high school, teachers provide a proficiency grade for the standards in which students have been assessed, using one of the following grades: ADV - Advanced, PRO - Proficient, DEV - Developing, NSP - Needs Support. Below is a description and visual of what these areas mean.

Advanced (ADV)	Proficient (PRO)	Developing (DEV)	Needs Support (NSP)
Advanced means that the student has advanced and in-depth understanding of the target expectations.	Proficient means that the student has successfully learned the skills required and can use them consistently without help.	Developing means that the student has basic knowledge, is working on the skills required and needs some help in order to be successful.	Needs Support means that the student needs continuous help in order to complete the skills.

Earning a grade report of "NONE" means that the student has not participated in class (whether in person or virtually) and has not submitted any assignments or assessments for the teacher to be able to determine a proficiency level. If this doesn't change, the grade would turn into an F at the end of the semester.

Beyond the proficiency scale, NHS administrators and members of the Targets-based Grading team communicated supporting rationale to families and students through letters and emails, school website, and student handbook changes. Additionally, other district policies and administrative rules related to grading were updated to align with the shift in grading practices.

Figure 4.4 *Communication to Families*

Report cards are, by definition, tools used to communicate student performance. The point and percentage systems of traditional report cards are familiar and comfortable to many of us. But simple points struggle to reflect the complex thinking and learning that happens in today's classroom.

Our new method of Grading for Learning (or standards based reporting) breaks down a student's grade into specific skills to communicate specific levels of achievement in each area. Previously, grades often reflected points taken off for behaviors such as tardiness, late assignments, sloppy work, lack of participation, or rude behaviors. Now, those are separated from our academic reporting and will be reflected in Employability Skills for Prompt and Prepared, Positive and Polite, and Persistent and Productive.

This school year, report cards will be sent home at the end of each semester. We share with you now a progress report of how your child is currently performing and is what the semester report card will look like.

Northeastern High School's district webpage offered the following rationale and justification for their recent shift and implementation of the Targets-based Grading system as seen in Figure 4.5 below:

Figure 4.5 *Northeastern School District Website Rationale for Change*

Why?

District believes the purpose of grading is to provide students and parents with a snapshot view of a student's proficiency on academic and behavioral skills at a particular time. The foundational beliefs of this system include:

- Academic grades are dependent upon student proficiency against identified learning targets and are not influenced by non-academic factors.
- Academic grades will reflect the student's level of proficiency attained on identified learning targets, as measured by specific assessments aligned to the learning targets.
- Behavioral skills will also be assessed and reported using defined rubrics:

NHS's grading rubric and subsequent communication to families aligned with the district's policy on grading. This policy was updated in July of 2020 to reflect the district's shift to a Targets-based Grading approach.

Figure 4.6 *Northeastern District Grading Policy*

5421 - **GRADING**

The Board recognizes its responsibility for providing a system of assessing student achievement that assists the student, teachers, and parents in understanding progress toward the learning goals and standards of the District's curriculum.

Grading and Assessment Policy

Academic achievement should be the primary factor in grades and is defined as the student's ability to exhibit progress in his/her ability to perform tasks, demonstrate skills, and apply knowledge to real world problems and situations.

Achievement can include subject-specific content, thinking and reasoning skills, as well as general communication skills.

The Board believes that the District's reporting structure should be a reliable system and one that ensures each student's "grade" accurately reflects his/her degree of progress in achievement of the identified learning goals and standards.

Progress of nonacademic factors such as effort, behavior, attendance, and late work shall be reported separately from academic factors. Extra credit will not be an option to improve a grade.

These documents were ongoing communications to families and other stakeholders as NHS worked through the transition process.

Study Participants

I conducted interviews with staff members following a semi-structured format that provided a foundation for questioning but also allowed opportunities for follow-up questioning as appropriate. Teachers were chosen through purposeful sampling with criteria that included at least three years of experience in the district teaching a core content subject area. Throughout my time at NHS I interviewed two teachers and an administrator (see Figure 4.7 below):

Figure 4.7 <i>Identification of Northeastern Teacher and Administrator Participants</i>			
<u>School</u>	<u>Name</u>	<u>Position</u>	<u>Experience</u>
NHS	Sarah Dean	English/Language Arts Teacher	7 years
NHS	Susan Smith	English/Language Arts Teacher	23 years
NHS	Sheila Larson	Former High School Principal; Current District Administrator	3 years 1 year as DA

The indepth student interviews also followed a semi-structured format that allowed follow-up questioning as appropriate. Students were chosen via purposeful sampling following criteria that students met a minimum grade point average and were intending to enroll in a college or university after high school. The list of student participants and their coded identities are as follows:

Figure 4.8 <i>Identification of student participants</i>		
<u>School</u>	<u>Female</u>	<u>Male</u>
NHS	NE3	NE1, NE2, NE4, NE5

The primary administrator who led the Targets-based Grading reform at NHS was former Principal and current District Administrator, Sheila Larson.

Administrator Perceptions

Larson spent three years in the district as the High School Principal and was the key administrator in implementing the Standards-based Grading reform at NHS. Prior to the 2022-2023 school year, she was selected as the new District Administrator. She made a point of noting that despite her shifting administrative positions, she still felt a strong, vested interest in continuing with the evolution of their Targets-based Grading implementation.

Administrator Background Experiences and Beliefs

Larson shared personal grading experiences that she had been subjected to that likely played a role in forming her early career beliefs. Anecdotal experiences from her time as a student prior to her educational career as a teacher and principal spoke to the familiarity and comfort that came with receiving ‘A’ grades:

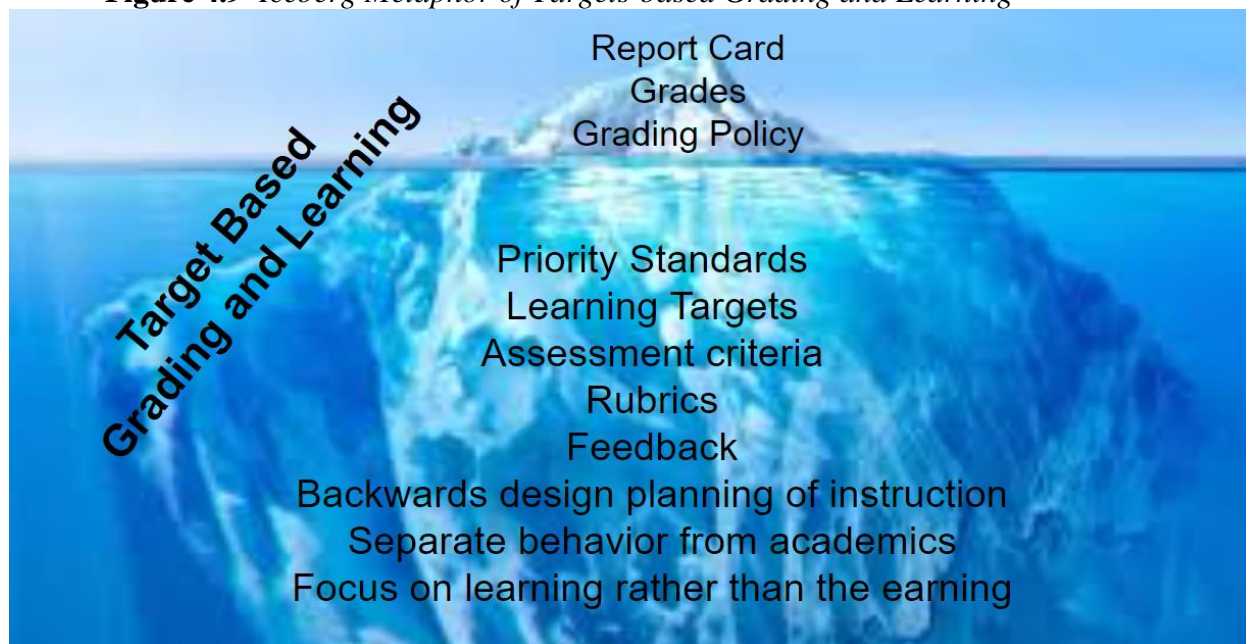
I think as a student I always strived to get A’s...straight A's. I worked very hard at that, but when I think back to that, I think there were times where I probably didn't really learn. I was just doing whatever I needed to do to get an ‘A’. And that would also be true in college as well.

Larson’s beliefs and perceptions about grading shifted and evolved throughout her experience and career, most notably as she grew as an administrator and began to recognize that traditional grading practices she experienced earlier in her life as a student and then as a teacher had much room for improvement. She understood that changing the practices of NHS from a traditional grading system into a Targets-based Grading system was going to take time, it was going to be challenging, and there needed to be much staff learning and communication.

Larson made a point to stress the front-loaded learning with staff and communication with all stakeholders that happened throughout the planning and implementation phases of their Targets-based Grading system. Prior to the start of the 2021-2022 school year and 2nd year of

Targets-based Grading at NHS, Larson shared the following graphic and message at her staff meeting highlighting the academic relevance behind the metaphor of the Targets-based Grading and Learning iceberg (figure 4.9 below):

Figure 4.9 *Iceberg Metaphor of Targets-based Grading and Learning*



Larson shared the following excerpt with her staff that accompanied the metaphoric iceberg image, which articulated NHS's intended academic relevancy and purpose behind the image:

We share a visual of an iceberg, because what is most visible is a grading policy and student grades and report cards will look different. What is less visible to you is the work that we have been doing the past two years—identifying priority standards, essential learning targets, developing assessment criteria for proficiency, rubrics that will allow for better feedback to be provided to students, backwards design planning instruction, separating behaviors from academic grades so students will receive a separate grade and feedback on their behavior. Our ultimate goal is that there will be a shift in students to focus on learning, rather than the earning of grades.

There were some questions at the policy committee meeting regarding the research behind this. The research on the impact on student learning as a result of Targets or Standards-based learning are immense...there's so much that we have read and continue to read in research studies, articles, books and dissertations but we know you don't have the time to read all that we continue to read on this topic.

For years, we have been studying the research behind what strategies have the highest impact on students and researcher John Hattie is well known in the education field for compiling the research and reports out the effectiveness in effect sizes.

Administrator Perceptions of Fidelity of Implementation

Despite all of the staff learning, communication, and time spent on developing targets and new grading criteria, Larson indicated there were still varying levels of implementation within the NHS building and frustration associated with a perceived unwillingness to better understand the Targets-based Grading system:

You know, we were at the point that we're fully implementing it, and then as principal, I would find that some people really weren't implementing it. They had their secret little grade book and were still doing it traditionally. And you know, kids were finding that out, parents were finding that out. And it became very frustrating from my end when I would get the parent phone calls or kids concerned about their grades. So I guess my most memorable thing is that it is challenging to switch to Targets-based Grading. It is a huge shift. I think it makes the most sense. It was frustrating for me when people did not want to learn why this could be better.

During the 2021-2022 school year, Larson shared the following slide during a staff meeting addressing concerns with the implementation as it related to staff practice and fidelity:

Figure 4.10 *Staff Meeting Slide Identifying Grading Inconsistencies*

We need to be consistent on:

- **ReAssessments** - allowing students to reassess, even to demonstrate advanced proficiency
- **Behavior** must be separated from academic grades (i.e. parent signature, group participation)
- **Transition to letter grade** - Advanced is NOT the only way to earn an A
- Identifying your Proficiency Criteria for learning targets

Larson shared this slide specifically as a means to help staff shift beliefs and better understand that successful implementation of the Targets-based Grading system required consistency of purposeful practices allowing students opportunities to reassess and separating student behavior from their grading outcomes.

Administrator Perceptions of Student Experiences

Larson indicated that many students were initially less than willing to do homework when the students realized it was not going to be graded; however, student perception shifted as students learned that despite homework being ungraded, it was a crucial component of the learning process.

So initially when we changed and said that they were not getting graded on their homework or formative work, then many kids then didn't do any of it. <laugh> Because what's the point? They're not getting graded on it. So teachers, when it came to the summative assessments and could point out that well, you haven't been doing the work, so this is why you were not proficient on the assessment. I think it took some time for kids to realize that formative work does matter to be doing that. Some teachers also had to shift their practices of what they do with that formative work to give more reason for

kids to do it. It's that transition from playing the "points game" of how much does this count for my grade versus focusing on the learning.

Larson also indicated that despite the initial reluctance of NHS students to do homework when they recognized it did not count, coupled by their ingrained behavior of focusing on points instead of learning, the pendulum has shifted somewhat this current school year as indicated by her response when asked about current student perception of Targets-based grading at NHS:

I hope that it has changed more to a focus on learning versus, you know, earning the points to get their grade, which is what it used to be. I know I've heard that, and this would be from teachers, so I guess I haven't heard this from students, but students seem to be more concerned about the feedback on the rubrics versus what is the grade that they got. I think it would be helpful to give a survey to students to get their feedback that way. I think their feedback is really important as teachers continue to work together as a staff on refining their practices. You know, I think there's some inconsistency in how reassessments are handled. That would be helpful for them to look at together.

Larson was fully committed and dedicated to the implementation of the Targets-based Grading system as she believed it would lead to better learning and an overall more positive experience for students; likewise, she was fully committed to supporting her teaching staff with their ongoing, student-facing implementation of the Targets-based Grading policies, practices, and school commitments.

Teacher Perceptions

Sarah Dean, a NHS English/Language Arts Teacher with seven years of experience, who identified herself as an advocate for the transition to the Standards-based Grading reform process that her school was immersed in. She presented herself in an extremely professional manner

evidenced by a very friendly and professional greeting towards me, as well as the informal questions and smalltalk as we first got to know each other. The passion for teaching and working with adolescents that Dean showed was obvious. She expressed more than once that she loved her job and loved working with her students at NHS.

Susan Smith, a veteran ELA teacher of 23 years, who indicated she had seen several initiatives throughout her career related to teaching, assessing, and most recently, grading. Smith was friendly, thoughtful, and was proud of her lengthy teaching tenure at NHS. She paused often throughout our conversations, giving thought and reflection to her responses. She shared the sentiment that she cared for her colleagues and her district administrator, Sheila Larson, while also sharing that the grading reform had sparked a wide emotional response throughout the entire high school staff, some of it negative and hurtful.

Teacher Background Experiences and Beliefs

Early on, Dean shared a personal experience from her earlier years as a student herself in a similar size high school prior to her career as an educator that fostered an early sense of comfort with traditional grading practices. She laughed while recalling the legacy practice of receiving extra credit as a means to pass her Chemistry class and to bolster her grade point average, while likewise recognizing that as a less-than-ideal contemporary grading practice:

I do recall though, passing Chemistry solely based on extra credit that was not content related. And that I think that is a capital offense <laugh>! I am so thankful that I passed Chemistry and I'm so thankful it didn't harm my GPA. But I think it's hard for me as a teacher to think about that poor practice that the teacher allowed us to get by on and to bring in class supplies and it would just bump up our grade. You know, at the time I think I felt like it was sort of unethical, but man, was I thankful as a sophomore! Because I did

not understand Chemistry. So that would be memorable, like a negative I guess you could say.

Dean shared another personal story related to traditional grading that has helped inform her current practices as a part of NHS's transitional process. However, as she was recounting this example of her experience in Spanish class, she spoke to the experience of how it feels to be a student who took a class and did not learn, and how that feeling helped shape her perspective as a teacher when it comes to her students, how she grades them, and how she hopes students perceive her execution of instruction and grading:

Another experience I had in high school: I looked forward to Spanish. I was so excited to take Spanish in seventh grade at the public school. And I was really excited to take eighth grade Spanish. And then when I moved to the private school, a small school with limited offerings, they didn't offer Spanish until sophomore year. I was devastated. I was like, okay, I can wait. And (when I did take it my sophomore year) it was a joke of a course. It was to get the credit. The teacher was, you know, proficient in Spanish, but he was also doing several other jobs. He was the coach and he was the IT guy and like just stretched too thin. And I look back at that too, like there were kids that shouldn't have passed that class. I worked hard because I actually wanted to learn a language, but I wasn't able to learn because of the experience that I had. And so I think that informs the way that I now grade and teach, because I want the students, when they leave my class to say, "I learned something with Ms. Dean." Last year when I had a couple kids tell me that my class was their hardest class, I almost got giddy. I was like, oh, wow, I'm doing something. This is great. I've had kids tell me before that my class was really easy, like it was practically impossible to fail my class. And I thought, you know, I don't want them to fail my class. I

want them to do well. But I also don't want them to look at my class as a joke or like a time filler.

Similar to Dean's experience, the experience Smith recalled was not entirely positive and this was an impactful experience that helped shaped her improved practices once she herself served the role of educator:

When I was a student, I honestly didn't pay too much attention to grading and how I was graded. I have a few teachers who stand out because I always had teachers who were very good about talking to you about how you were doing and stuff. I was one of the lucky ones. I think that all through school, even way back when, I had an English teacher who stands out because every essay that you turned in, it was the holistic grading, you know, like ABC or whatever. He would fill the back of each paper with just comments and discussion about what you had written about. And that is what stood out to me because it was a joke. His handwriting was terrible. So like the family would sit around and try to figure out what <teacher name> had written about. But the fact that he took that kind of time really stood out that it didn't matter what grade was ultimately on there because he had, you know, he cared enough to comment on it.

Dean described the origins of her grading practices as rooted in her experiences as a student herself and her teacher training program. She spoke passionately as she explained inadequacies of her teacher training program and compared her experiences using traditional grading to the reformed Targets-based Grading system:

Partly because of my inexperience as a teacher, I think I had no direction when I began. I didn't know how to grade students. Targets-based Grading gave me a framework to work within. I think I could be successful as a teacher and my students could learn in a

traditional system as well. As I said before, you know, it's the people, not the programs, right? And so I think if I learned how to properly implement, you know, high quality rubrics and, and give effective feedback and give students many chances to like show their knowledge and not just like once and done, um, which I think I would've figured out over time. But this gave me a path forward and I, and I saw grading and teaching through a different lens. I think that it's the kind of lens that teachers, effective teachers probably were already working with.

Both Smith and Dean shared the sentiment that their personal experiences from their past as students were variables that affected their perceptions and beliefs regarding grading in their current roles as high school teachers.

Teacher Perceptions of Student Frustration

Both Dean and Smith recognized that their building administrator, Sheila Larson, worked hard to craft staff learning and professional development around their implementation of Targets-based Grading, and worked hard to change the belief systems staff may have had related to traditional grading practices. Nonetheless, Dean also indicated that students have expressed frustration with a lack of clarity between the new Targets-based Grading and how that corresponds to their more familiar, more traditional letter-based grading experiences.

I mean, it's understandable why the kids get so bent out of shape about their GPAs being affected. That was, you know, one of the other things I've seen in some of the student feedback I wrote down that they really, they're really not clear which letters correlate to which scores. Because some teachers mark proficient as an 'A' and some teachers mark proficient as a "B" and then at the end of the day, the kids want to just convert it right back into the "ABC," you know. They said the system's really screwing up their grades.

They shouldn't even have a letter grade if we're gonna do this with efficacy. You know, they, they don't want their grades falling right here at the end of their high school because of inconsistency with the grading system and teachers. And that doesn't allow them to apply for colleges that they, you know, that they've been wanting to apply for. Like how, how can they be successful?

From Dean's perspective, much of the student frustration is rooted in a lack of implementation consistency among staff. Smith likewise indicated that there was frustration from students throughout the transition of grading practices to the Standards-based System:

There was a lot of frustration. I remember some of our top students, not so much that they weren't succeeding anymore. They were still doing excellent, but they felt they understood less of what was expected of them. And there's only so many times you can say, "keep doing what you've been doing and you're going to be just as successful." Some of the lower students got frustrated because even though, in theory, the rubrics were telling them specifically what they needed to do, not everyone had those rubrics yet. Or if they had the rubrics, they weren't very useful yet. So they weren't sure. "How am I going to pass if I'm not sure how to do that?"

Smith, who is also a parent of a high school student at NHS in addition to being a teacher in the building, articulated some confusion from her son's as well as a parent perspective, and suggested that the overall community has still not bought in to the Targets-based Grading reform at Northeastern:

My son's a student here. So I see the parent end and the teacher end. And when I can't necessarily explain what his grade is reflecting using the targets for some of his classes, I feel like I can only imagine where parents outside of it and the students need it translated

all the time to them. Even though I give them the rubrics, we talk about it beforehand, they still won't have that understanding of, well, "What's my grade? Like, what percentage is it? How many questions can I get wrong?" Then there's the rest of the community: the buy-in isn't there yet.

It was obvious that Smith is a committed, dedicated teacher who, despite acknowledging some staff and student concerns with the implementation and timing of the Targets-based Learning system at NHS, was supportive of the process and of her administration.

Teacher Perceptions of Overall Implementation of Targets-based Grading

Dean shared further thoughts about overall implementation and her perspective on progress of implementation related to this current school year. It was apparent that Dean was herself a bit frustrated with the execution she and her colleagues had provided in the grading reformation process, but she also indicated that there were signs of overall growth and improvement. She proudly stated that she believed in this process while also recognizing that some of her colleagues did not share her positive sentiment.

I think that this year, our implementation has been, well still not good, but more consistent in the expectations from admin on how we grade. And, rather, how we put things in the grade book. That's what's really clear in standards. And we don't necessarily have a lot of accountability in the other areas because administration are strapped and they don't have the time to sit down with all of the departments and ask us what kind of grading we're doing in our classes, or at least they're not doing those things. And so then that leads me to believe like that I can do what I want in my classroom. Now I really believe in this process, so I'm striving to do it with efficacy, but I know other staff don't.

Dean shared her perspective that despite the efforts of the two years of preparation, staff training and development, and district communication, some traditional practices were still prevalent at NHS even after the transition to the Targets-based Grading approach, suggesting that the implementation was not being executed with fidelity, and may be a primary source of student confusion and frustration. Dean presented herself as a passionate educator who is deeply invested in the culture of NHS as well as the success of her students and overall execution of the Targets-based Grading system they currently have in place.

Although supportive of the district and administration, Smith also recognized some concerns and issues with the Targets-based Grading implementation process that Northeastern had undertaken. She identified pushback from collegial staff at NHS related to implementation of the reformed grading system:

I saw a lot of pushback, which concerned me because I knew it was going to hurt all of us. And by trying to go along with it, it would hurt me because this teacher's not doing that. So there was a lot of concern about it. And I mean, some of it I'm seeing now just because there were some groups that even though we're technically in our second full year of it, they're still finishing their learning targets and don't know what their rubrics are. And it's not their fault. It's just the way life happened with COVID and all that.

The pushback that Smith recognized was consistent with variables that Principal Larson identified which in turn became impetus for ongoing professional development opportunities for staff. The teachers' pre-existing biases and belief systems were a barrier in the implementation process which prompted an engaging discussion at a staff meeting in the 2021-2022 school year as seen in figure 4.11 below to address staff member beliefs and biases related to their implementation reluctance:

Figure 4.11 *NHS Staff Meeting Prompts Addressing Staff Pushback*

Do you agree/disagree with these statements?

1. Grades should be the basis to differentiate students.
2. Grade distribution should resemble a normal bell-shaped curve.
3. Grades should be based on students' standings among classmates.
4. Poor grades prompt students to try harder.

Smith expressed some disappointment and discontent with her colleagues' resentment towards the philosophical grading reform; likewise, she was disappointed with the lack of initiative shown by some colleagues as it related to timely completion of learning targets and subsequent rubrics. She also indicated that the implementation of the grading reform process may have been rushed which subsequently affected the overall fidelity of the intended Standards-based grading system:

I honestly think the biggest mistake we made as a district was we kept to our timeline even though COVID hit. So when we first started implementing it, it was like, from this school year, these are the things we're gonna do this school year. And we came back from going online in COVID without having done that last half, but still started it. I would've said pause, we need to take a break.

There was obvious passion, dedication, and commitment to students expressed by both Smith and Dean throughout all our conversations. After hearing various perceptions from staff regarding students' reception to the change in grading practices, I was eager to learn from the students themselves.

Student Perceptions

I interviewed five NHS 12th grade students through a predetermined schedule that had been communicated to each of them prior to my arrival. The student interviews took place in a moderately-sized multipurpose conference-like room that had an entrance connected to the office and a second entrance directly connected to the hallway near several other classrooms. Students entered independently via either entrance; it was obvious they were no strangers to this multi-purpose room which I later learned served as a meeting and conference room for various clubs, activities, and other miscellaneous student or staff gatherings. The Northeastern students were friendly, talkative, and respectful. The students expressed an appreciation for the opportunity to talk and share their thoughts about grading and their school experiences.

Student Comfort and Preferences with Grading

Students shared various perceptions, beliefs, and anecdotal experiences related to their preferences and comfort of both grading systems. One student in particular spoke to their personal perceptions of assessments and retakes which reflected indifference between the systems and suggested that despite the change in grading, instruction and assessment practices have largely remained the same.

I feel like not a whole lot has changed, assessment wise. I feel like a lot of the teachers give the same type of assessments and the same type of assignments. I think just the way our grades are being recorded is different. At this point that's where we're at. I think that some teachers do it a little differently than others, so some have changed their assessment and their retake policy to conform a little bit more with the standards and the like, because the point of standards is, you learn the content, not necessarily just getting the question right. And I think some teachers do that a little bit better than others. (NE1)

Two students indicated that the previous system entailed less actual emphasis on learning and had more emphasis on a short term memorization-based system that required students to learn snippets of information for several week periods of time and then regurgitate knowledge and understanding of those specific chunks of information on a follow-up assessment.

I like Standards-based because I like, again, because I think that the goal of Standards-based is for you to understand the content. Whereas like personally when when we were in the other grading system, like I got good grades, but I don't think I ever really retained as much information as I do now because I think Standards-based focus is more on you knowing exactly what you need and like you working towards that and you being able to apply the different skills in classes. Whereas the old grading system it was, it was basically just like a memorization thing, like where you, you know it for the test, but then after the test doesn't really matter because you don't have to apply it anymore. (NE2)

It appeared that a couple Northeastern students appreciated the concrete, dichotomous nature of the previous system that indicated whether students either passed or failed, there was a finality to the process. Comparatively, the more descriptive, likely more vague indicators of the new Targets-based Grading system with a middle ground of “approaching” suggests additional learning needs to occur for students to reach proficiency.

I definitely think the ABC system that we always had was easy to navigate because, I mean, you either passed or you did decent or you failed, you know, like it's just simple. And now when we are looking at all these different categories that there are, like your employability skills, just all the different subcategories there are, it's a lot more intricate. So there's more data, but it's still not as, it's not simple. (NE4)

Students had varying and inconsistent perceptions of ease and simplicity as it related to both the traditional and Targets-based Grading systems.

Student Perceptions of Grading Practices Related to Post-Secondary Opportunities

Multiple Northeastern students associated their future prospective college plans with their thoughts and beliefs about the current and past grading systems. Students were concerned about their upcoming college experiences and it was apparent that students recognized there was not going to be a consistent, standardized grading system across the colleges they considered attending next school year. The implication and direct mention of college was prevalent in almost every Northeastern student dialogue. One student also offered a perspective related to academic relevance that included thoughts related to homework, college admissions, and other variables related to improved learning:

I know that a lot of colleges, like I know somebody who works, who used to work in admissions at a college, and he always said that rather than looking at transcripts and GPAs, he always looked at like writing and why you wanna get into the college and like the essays and more of like the personal aspects. So I think that depending on the college, it really won't hinder someone's ability to get in. (NE1)

Several students recognized that traditional practices within the new grading system as it related to a traditional grade point average were still in place contrary to the grading system reform approach. Students believed that GPA was likely still relevant despite the transition of grading systems:

So what's the point of um, like going through all this work to have Standards-based, but then ultimately it just gets turned right back into a letter grade and GPA? I'm not sure if with Standards-based—like we're still fairly new in implementing it—I'm not sure if a GPA

is like supposed to be within Standards-based or not, but I feel like a GPA is fairly important for colleges to look at. (NE5)

All students saw a connection between the reformed grading system they were experiencing as it related to their college plans their next year following graduation at Northeastern.

Student Perceived Value of Targets-based Grading

Several students spoke to their school experience within the reformed grading system. Students recognized the academic value in the use of a Targets-based Grading approach and also identified the role that rote memorization has played in traditional grading. Most students also indicated that although they may not necessarily prefer the reformed grading systems, they recognized that the Targets-based system generally had a greater impact on learning.

I think the benefits, if Standards-based is implemented correctly, is that your students are learning more about the content and they show that they're improving based on like learning, rather than the old grading system that we use. It's (the old grading system) kind of just like, you memorize the stuff, write it out, and then you forget it after the test. But Standards-based really focuses more on do you really understand the content and can you explain it and can you apply it. I think specifically for our school, the challenges that, because we tried to implement it before COVID, I don't think it's necessarily being implemented to its fullest. Like I think that there's still parts of our old grading system that are merging with the Standards-based, and I think that's creating a problem. But I also think that it's probably a problem because like not everybody in the nation is changing to Standards-based grading. (NE1)

Several students recognized value in the Targets-based Grading, particularly a wider opportunity to demonstrate proficiency and learning.

Student Perceptions of Multiple Opportunities for Assessment

Two of the Northeastern students recognized and appreciated that one of the primary tenets of their reformed Targets-based Grading system revolved around the notion that students should be given multiple opportunities to demonstrate their learning. Although some of the students seemed to emphasize grading outcomes over learning as the primary motivator, some of them assuredly recognized the value of the actual learning that was likely to occur given multiple chances for student to show their mastery:

I definitely think Standards-based grading is more set up for that because the way that it's supposed to work is that, I mean, at least the way I understand it, that you can basically reassess as many times as you want, as long as you are improving every time. And that, that allows you to get all the way up to Advanced because it's more focused on you learning the content rather than you just getting good grades. (NE1)

Not every student's perception of having multiple opportunities to show their learning and proficiency of targets was positive. In fact, a couple students indicated that there were many students at Northeastern who were taking advantage of the multiple retakes to game the system without putting forth an initially strong effort to learn and understand the material. One student, for example, called out his peers and was critical of the Targets-based Grading system by suggesting that it actually created less responsible, less initially robust academic habits in students:

The retake policy is different so that that could be a benefit. You know, you get pretty much unlimited retakes up to full credit back, so that's nice. I kind of think that the kids that kind of blow it off the first time you take the test and just take it, then getting unlimited retakes kind of strengthens that mindset. It's like, oh, I can just retake it again.

Oh, I failed it another time, it's not a worry. I can retake it again. It doesn't necessarily enforce good studying habits or deadlines or response to deadlines, I guess. (NE2)

Despite the lack of consensus from students on multiple opportunities for assessment, it was apparent to me that all students understood the connection to improved learning and grading outcomes.

Student Confusion with Targets-based Grading

Confusion and issues with communication related to the Targets-based Grading system were prominent themes among the Northeastern students. Some students liked specific percentages because they perceived it to be less confusing as it pertained to how well they were performing. For example, one student indicated that seeing a specific percentage that is borderline on both passing and failing (e.g., 60%) as a specific number is more alarming and motivating for improvement than receiving the less precise, narrative equivalent (e.g., Developing):

I see it as like if you got a certain percentage on a test or something that, you know, it was, if you got like a 60% on the test, like you were, most people at least I, I would, I would immediately communicate and say, no, I didn't do good on that, I want to redo it. So I guess in a sense I'd say the old system is, is more alarming, you know, if you do bad on something. But, uh, the new one it is, it's more, uh, I'd say it's almost a hidden in a sense because, uh, like trying to find the, like the score I got on that test or something, it's a lot harder if I have to look through all the different categories and like add 'em up or, and like put them into one number or one, one of the four like advanced proficient, like it's, there's just so much there that it's, it's harder to, uh, you could fine tune like certain things that you did bad on. Like, uh, I had an English test where I did really, really good

on the, um, like the grammar part of it, but when it came to the actual, uh, like the con like the actual meat of it, like, like it wasn't quite, I got worse, a worse score on that, so I could just fix that instead of the grammar. Okay. So that would make it easier, but it's also, it's easier just to see that you did bad with a simpler system. (NE4)

Similarly, another student voiced confusion related to the Targets-based Grading descriptive language and indicated they were more comfortable with the previous numerical system of traditional grading.

I think that the older system was a little bit easier for kids to understand because it was all numerical and it's really easy to track because Standards-based, the way that it works here is a little bit more subjective and it changes between classes, so it's kind of harder to follow unless your teacher is really able to like explain that every time you get a new assessment. (NE1)

Despite the reformed grading system being more descriptive in student performance related to learning targets, confusion about precisely how well students were performing was brought up more than once. One student shared that re-takes and re-assessment opportunities existed in the previous system as well as within the Targets-based system but the means of doing so was much less confusing within the previous, traditional grading system.

I would say the old grading system was easier to understand and it seems more applicable. Like for those students that are going into college after high school, Standards-based seems kind of soft and very forgiving towards individuals that don't necessarily apply themselves all the time. You could still change or improve your grades with the old grading system. Like you could do a retake, it was a little bit harder. A lot of teachers had like a form you had to fill out. So it was just a little bit more time consuming

to do, but it was still possible. And now with this Standards-based, some teachers still have forms you have to fill out, but for the most part it's just, "Hey, I want to retake this" and you're allowed to retake. I would prefer the old system. It seems like a lot, it seems less confusing. (NE4)

Students expressed that although the new reassessment process was confusing and inconsistent between teachers, they liked the additional opportunities for assessment and subsequent grade changes.

Student Perceptions of Teacher Change

Throughout the interviews, several of the students expressed that they felt the teachers were more comfortable with the traditional grading systems and although the grading and reporting system had changed, comparable changes to instructional or assessment practices remained relatively unchanged. Instead of an evolution in an overall learning system, the suggestion was that a lot of the same practices were still in place. Multiple students suggested that some teachers had not yet entirely bought into the new system and perhaps even a small handful of teachers were downright resistant. One student indicated that the teacher changes with the new system were superficial at best:

Like I don't know. I think teachers just, it just feels like they give us busy work and nothing has really changed with the new grading. I mean teachers had to change their grading but stuff has kinda changed, and kinda not, too. Not all teachers have changed really. With the old system the grading was more open and with the new, the new one is more hidden. The old system had more scores and points, like data, a lot more data available so we just kinda knew. The new grading feels like hidden and less consistent. But it's more real world. (NE3)

Similar to the multiple student perceptions of teacher resistance, one student indicated that one particular teacher was still using the previous, traditional system within their current reformed system:

I'd say a lot of teachers still use, like their old grading system. Like, I mean, Mr. (teacher name), he was like one of the old style teachers in a sense, you know, like he used where either you passed, you did decent or you failed type of thing. And there's still a couple teachers in the high school who do that. But, it also helps. I feel like they're a lot more prepared to help you grow from that if you did bad. (NE4)

It was evident to me throughout the Northeastern student conversations that they perceived a wide range of teacher implementation of the Targets-based Grading.

Student Perceptions of Teacher Inconsistency

Student perceptions of inconsistency among individual teachers came up often as it related to Northeastern's narrative descriptors of grading. Students recognized teacher autonomy throughout the NHS faculty based on the inconsistency of the Targets-based Grading execution between their different courses. Students indicated that individual autonomous teacher practices varied from department to department, teacher to teacher, classroom to classroom as it pertained to Targets-based Grading in general. Students also expressed that there was teacher inconsistency within the delivery of Targets-based grading practices, specifically the differences and friction between "Proficient" and "Advanced." One student highlighted this specific lack of consistency and uniformity:

I guess this is more on like the implementation side of it, but implementing a new grading system, like I understand that's a lot of work especially for an entire district, but I feel like having more meetings, kind of planning it out more to make it go smoother and get

everyone on the same page would improve it a lot. Right now we have, like I said before, there's some teachers where “Advanced” means one thing and other teachers where “Advanced” means something completely different. And so, a lot of inconsistency or lack of uniformity between teachers. (NE2)

Some students indicated beliefs that teacher confusion with the new Targets-based Grading may be a cause behind the teachers’ inconsistent applications of the two indicators (“Proficient” and “Advanced”).

The NHS students were not shy, and they were open and honest when sharing their voices regarding their experiences and perceptions of their school’s transition to Targets-based Learning.

Case Two: Southeastern High School

School Context

Visits to Southeastern High School (SHS) began in December 2022. SHS is a small midwestern high school located in a rural community of approximately 3,300 people. The entire 4K-12 district consists of three buildings all on a connected district campus property. SHS is a rural midwestern school of 320 students of which approximately 90% are white and approximately 10% of students are Black, Indigenous, and people of color. Approximately 42% of students are economically disadvantaged. SHS demographics and enrollment have largely been consistent over the past five years although they have seen an increase of students in poverty (34.2% to 42.4%).

Figure 4.12 *Demographics for Southeastern High School*

Figure 4.12	
Southeastern High School demographics	
Enrollment	323
Percent Open Enrollment	8.4%
Students Economically Disadvantaged	42.4%
Students who are English Learners	1.2%
Students with Disabilities	10.2%
Students who are White	89.5%
Students who are BIPOC	10.5%

As evidenced by the state level education department's school report summary, SHS's self-description is that it prides itself in its dedication to helping all students achieve high standards both in the classroom and beyond the classroom. To that end, SHS also indicates it provides a challenging curriculum to a wide range of students and student interests.

Southeastern High School Community

Upon my arrival into the Southeastern city limits, it was immediately evident that this was a community that valued its students' accomplishments at SHS. I noticed multiple signs of athletics accomplishments including recognition of a fairly recent State Championship in softball. As I journeyed further into town en route to the high school building, I saw further evidence of community support in the form of large community business advertisements on signs lining the outfield walls of the local ball diamond. When I entered the High School building (doors were unlocked due to students entering at the start of the school day), my attention was immediately drawn to massive portraits and plaques lining the initial walls of the entrance and commons area: the SHS Wall of Fame. Past alumni, shown in full color photographs, were

recognized for outstanding accomplishments in athletics, coaching, and co-curriculars. It was clear that the placement of the Wall of Fame placards as the first noticeable detail upon entering the premises was intentional.

The school pride of SHS was evident throughout the building. Multiple students meandering through the commons area prior to the start of the school day were wearing their signature colored letter jackets. Students were talkative and full of energy for early morning. As I wandered through the large vestibule and commons area into the high school office, I noticed immediately that one of the high school office staff wore a SHS sweatshirt—yet another sign of a proud community and proud school culture. The demographics of the early morning school community appeared commensurate with that of the small, midwestern town and school.

SHS Transition to Targets-based Grading

Southeastern High School's initial planning for their transition to Targets-based Grading occurred primarily in the 2019–2020 school year in the form of several professional development days and monthly after-school staff meetings throughout the entirety of the school year. SHS defined their Targets-based Grading model based on the work of Hierck and Larson (2018) as a system of grading that requires teachers to grade and report against specific learning targets, and not standards as a whole. This grading model allows parents and students to clearly understand what children are expected to learn in class, and also provides them feedback on progress toward specific learning targets.

The majority of this professional learning and monthly meeting time was used for individual teacher creation and revision of learning targets, and alignment of success criteria to the specific targets. Some small department work was done as a team but given the limited number of staff in each content area, SHS administration chose to focus the majority of their

attention on individual targets. Other than the initial justification of the rationale, there was little time spent during the 2019–2020 planning time on staff perceptions, beliefs, and understanding, or challenging staff counterpoints to the adoption and implementation of the Targets-based Grading system.

Southeastern’s grading transition journey largely became implementation reality during the 2020–2021 school year. The professional development time in 2019–2020 resulted in their current grading reform system that utilizes a Targets-based Grading approach ranging on a zero to three scale, from a zero (No Evidence) to one (Needs Support; Not Yet) to two (Approaching; Almost Got It) to three (Mastered; You Got It). Figure 4.13 below denotes the proficiency range for SHS Targets-based Grading criteria:

Figure 4.13 *Proficiency Criteria for Southeastern High School Targets-based Grading*

3. Proficiency Criteria Teachers will use district-approved proficiency criteria. Teachers may create additional proficiency indicators for each target in conjunction with the district approved proficiency criteria.	Mastered (You Got It) 3	Approaching (Almost Got it) 2	Needs Support (Not Yet) 1	No Evidence 0
	The student demonstrates mastery of the learning target.	The student is progressing and shows partial mastery of the learning target.	The student has gaps in mastery of the learning target. Guidance and practice are necessary in order to progress.	The student has not submitted evidence to justify progress toward mastery of the learning target.
	Additional Explicit Descriptors	Additional Explicit Descriptors	Additional Explicit Descriptors	Additional Explicit Descriptors

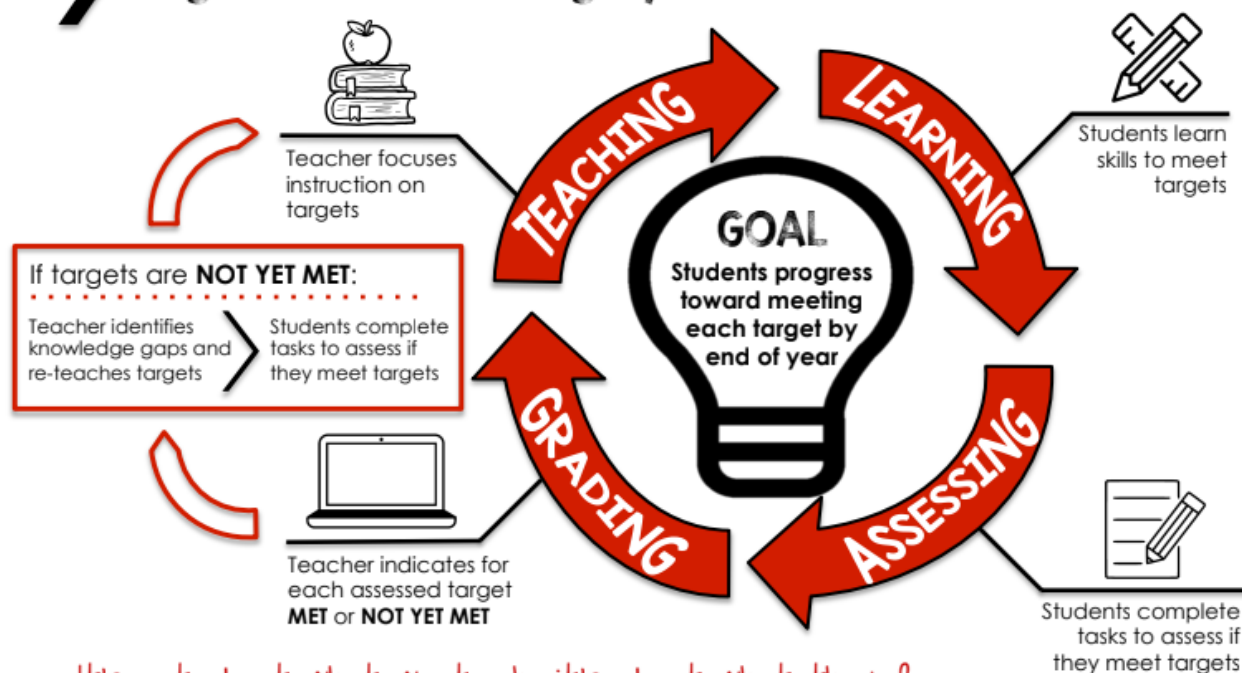
Since prior to the start of the 2020–2021 school year, SHS has delivered intentional communication both defining their approach to Targets-based Grading including their operational definition of Targets-based Learning, their rationale for grading using targets, and their cycle of learning explaining the role of the grading process in their Targets-based Learning approach (Figure 4.14 below):

Figure 4.14 *Southeastern High School Rationale and Operations of Targets*

What is TBL?

Target based learning (or TBL) is designed to only reflect true evidence of learning. It provides a clear picture of what the student has or hasn't mastered without the influence of other factors, such as effort and attitude. Consistency can be ensured with a teacher-provided rubric that establishes clear expectations and explains exactly what the student will need to master.

Target Based Learning Cycle



- It's not about what we teach, it's about what they learn.

Why?

Grading using targets accurately communicates achievement of learning targets to students, parents and educators. Grades are based on how well a student comprehends a specific target and not a percentage of work completed. Target-based proficiency grading empowers students to understand their own learning strengths and areas that need improvement. Students get direct and timely feedback on the areas/skills that they need to focus on.

Southeastern administration believed that their rationale and operations graphic not only thoroughly explained what the Targets-based Learning and grading approach entailed, but it also served to help all stakeholders understand why the district was transitioning to this grading system. Special emphasis was put on the fact that the foundation of this system had more to do with student learning than it had to do with what teachers were teaching. This graphic also

strongly aligned with their newly implemented practice of allowing and encouraging students having multiple opportunities to assess and reassess prior to a finalized proficiency grade.

Study Participants

I conducted interviews with staff members following a semi-structured format that provided a foundation for questioning but also allowed opportunities for follow-up questioning as appropriate. Teachers were chosen through purposeful sampling with criteria that included at least three years of experience in the district teaching a core content subject area. Throughout my time at NHS I interviewed two teachers and an administrator (see Figure 4.15 below):

Figure 4.15 <i>Identification of Teacher and Administrator Participants</i>			
<u>School</u>	<u>Name</u>	<u>Position</u>	<u>Experience</u>
SHS	Louise Clark	ELA teacher and Instructional Coach	6 years
SHS	Kevin Karls	Math and Chemistry Teacher	27 years
SHS	George Greggerson	High School Principal	10 years

The in depth student interviews also followed a similar format. Students were chosen via purposeful sampling following criteria that students met a minimum grade point average and were intending to enroll in a college or university after high school. The list of student participants and their coded identities is as follows:

Figure 4.16 <i>Identification of Student Participants</i>		
<u>School</u>	<u>Female</u>	<u>Male</u>
SHS	SE1, SE4, SE6	SE2, SE3, SE5

The key staff member behind the implementation processes and procedures was their high school principal, George Greggerson.

Administrator Perceptions

After a warm office greeting and meeting the district superintendent as well as High School Principal Greggerson, I was escorted to a quiet office wing that housed business offices, student services, and related services offices. I was given my own office area for all the interviews and was assured that students are more than comfortable coming into this area of the building. My initial interview at SHS was with their primary administrative visionary of the grading reform, Principal Greggerson.

Administrator Background Experiences and Perceptions

Not only is Greggerson the sole administrator at SHS, serving his 10th year in that position, he was a former teacher within the building 17 years prior and also served as Athletics Director for nine of those years. Throughout his 27 year tenure at Southeastern, he has coached football for all 27 years and has been successful in doing so, leading his team to the State tournament on multiple occasions. Greggerson spoke of school, district, and community pride. He has enjoyed his career at SHS and has proudly brought up his two children in the community and through the school system. Throughout our conversation, he received multiple phone calls, text messages, and calls on his building radio. It was obvious that Greggerson was an integral part of the day-to-day building operations. He suggested that he spent a lot of his day “in the trenches” directly working with and interacting with staff and students. It was apparent that he is deeply invested in the school, school culture, the overall community, and the students and staff he serves.

Administrator Perceptions of Implementation

Greggerson expressed he was passionate about the Targets-based Grading reform and the impact it has and will continue to have on students. He also spoke openly and honestly to the initial difficulties related to implementation of the Targets-based Learning/Grading system:

It's been a journey. There's a lot of, a lot of, it was tough for the first couple years. It was really tough. I think we, knowing what I know now, I would've done things much differently. We're kind of doing things backwards. I would've spent more time on changing our culture and our building and things like that and getting them on board slowly through the culture rather than just hitting 'em full board with this is what we're doing, this is where we're going. Because there was a lot of resistance and a lot of, and there still continues to be a little bit, but it's not nearly what it was. It's a lot of sales. It was like being a car salesman, you know, probably not even a new car salesman but a used car salesman trying to sell it, trying to stay a step ahead of everybody.

Greggerson also talked about the building culture during implementation. Although he is a servant leader and longtime colleague whose tenure in the building predates most of his staff, he recognized some of the staff changes that ultimately had to happen before a critical mass was achieved and progress was made with implementation:

I don't feel like I probably was as prepared as I could or should have been, but I don't know how much more prepared I could have been. Yeah, I think that, I think the culture and, and there's some, you know, some, some tier one issues and, and culture issues that I would've definitely solved before we got to Targets-based Grading because it's pretty tough to have a, you know, we have had a pretty divisive culture but it's starting to get quite a bit better. But it just created divisiveness in our staff and there was pockets of unhappy, angry, and another pocket of, "yeah, let's just go try this." And it just was not

cohesive. But it's gotten to that point. Some of it is from attrition too. I won't lie. Some people had to make the decision to stay or go and, you know, I've gotten to be pretty good friends with the principal up at (another school district) and he's been doing Targets-based for a long time and he said the same thing, that some people just had to get off the bus before he was gonna be able to move.

Greggerson indicated that he believed it was inevitable that some staff would not be on board, but he was not apologetic about the notion of challenging his staff to change, grow, and improve. He indicated that he personally felt stress throughout the implementation process. Greggerson expressed that the district's approach to implementation of the Targets-based Grading system was wrought with deep, deliberate planning and intentionality behind the district website and guiding documents communicated to all stakeholders despite earlier suggesting that had he done it all over again, he would have moved even slower and worked on bolstering staff culture prior to changing the grading system.

Administrator Perceptions of Student Experiences

Throughout my time at SHS and time spent with Greggerson, it was apparent that his passion for the transition to Targets-based Grading came from a desire to have a positive impact on students. As such, it was not surprising when he indicated that he and his staff took feedback from students during the initial implementation process of their Targets-based Grading system:

But I also, I also believe that you do your best work when you get challenged and you do get frustrated, you know what I mean? I think you improve and you get better when you are forced to change. So, you know, I would definitely do some things different for my own stress level, but I don't know how much, you know. Some of the people, you know how it is sometimes when you're working with some people, they are never gonna buy in

no matter what you do. So did students play any role in the transition process? Yes. But probably not to the extent that I would've liked to have seen them involved, but I'm starting to see where we are. Yeah, we took some feedback from students; probably not as much as we could have.

Not only did he acknowledge that it was important that SHS receive student feedback prior to implementation, he indicated that a primary strategy his school used to slowly get students and staff more comfortable and acclimated to the change in grading systems was via a pilot group during the planning year prior to full school implementation:

We tried to do a pilot group, kind of a pilot group with one teacher and pilot it and try and work through it. We had some pretty good success with it. I think that allowed us to be able to tweak some things and improve some things, and make them better. Before we implemented it, every teacher had to pilot one class. It came down to communication, kids understanding the system and knowing it, and now we have an entire system that wraps around it. We even have a built in reassessment period during our day. So kids have time to be able to do that.

Greggerson's thoughtful, matter-of-fact communication style throughout our conversation appeared consistent with the deliberate approach SHS appeared to take throughout the initial implementation process and first two years of execution of the Targets-based Grading approach. Greggerson also recognized that along with good administrative execution of the grading reform process, he needed buy-in and commitment from his teaching staff.

Teacher Perceptions

The first teaching staff member I met with and interviewed at SHS was Louise Clark, an English/Language Arts Teacher and Instructional Coach. It was obvious immediately that Clark

was a positive, friendly, high energy teacher who valued students and valued Southeastern's transition to the Targets-based Grading system. Clark was an unapologetic advocate for the Targets-based Grading transition and she expressed that she felt she had an obligation to be a building champion for the transition.

The second teacher interview was with Kevin Karls, a 25 year veteran Math and Science teacher who wore an entire wardrobe of SHS school spirit: a coaching pullover jacket, track pants, and tennis shoes. Karls spoke at length about the positive school climate, excellent student behavior, and overall level of student involvement in school athletics and co-curriculars. He spoke passionately about connecting with students on a personal level, the impact of student relationships, and the value of student involvement in co-curriculars as a lever to keep teaching and learning relevant to those students. He was also passionate about policy and systems changes in education—including their recent shift in grading practices—needing to be relevant to students.

Teacher Background Experiences and Beliefs

Clark indicated that she had overall positive experiences throughout her life as a student as it pertained to school, learning, and grading. She shared some of her personal experiences with traditional grading from her years as a student impacting her ability to reassess and subsequently achieve a higher grade:

I never got to reassess on things. I remember when I graduated. I was fourth in my class or really third in my class because I think the two in second place were tied. And I remember having a lot of resentment about that because I had a full schedule. I didn't have study halls. I took band and choir, I took all these things that those three guys who were my friends and everything didn't take. And they kind of purposely gave themselves

a lighter load and then finished. I think I got one 'A-' in high school and that caused me to be fourth in my class and not get to be Salutatorian with them or Valedictorian with them. And that was like so painful. Like I worked so hard for it and that teacher who gave me that 'A-', it was like one book report thing. And if I had had the opportunity to redo it, I would have. So it's just so funny looking back now on what that could have, what that could've meant. But also, like, it doesn't really matter that I wasn't Salutatorian.

Clark recollected her story about missing out on the Salutatorian award with emotion and disappointment. As we transitioned our conversation from her experiences as a student to her experiences as an educator, Clark also spoke to her perceptions of validity with Targets-Based grading as it pertained to her experiences as an English teacher at SHS:

I remember doing a lot of rubric writing from the English teaching perspective of like how to write a good rubric with a point by point breakdown of the skill. And that actually translates kind of well to Targets-Based learning in a lot of ways. But then it's not the same. And some parts of that drove me a little crazy when we were starting Targets-based learning. I'm like, "But it's a whole essay!" There's many things that I'm grading them on. How can it, they only get one grade in the end or a couple target grades in the end? And that was a really hard thing to break my mind free of because my instinct was like, I, I can be more specific with this rubric, the traditional way down to a point than I can be with these three targets that are just gonna be a three, two, or one.

Both Clark and Karls spoke positively of their formative and high school years in education. It was evident that they both believed the academic and grading experiences they had personally experienced were generally positive as it related to their eventual career paths as public education high school teachers. They both acknowledged that high school public education

should be evolving and should be looking to improve grading outcomes and experiences for all students.

Teacher Perceptions of Student Experiences

Throughout our conversation, both Karls and Clark were open and honest about some of their building's implementation hiccups. Clark spoke to an initial negativity, discomfort, and confusion from students related to the transitional process:

There was a lot of negativity about it at first. Yeah, they just didn't understand it at all. I don't know how well we explained it to students. I think that was left to every individual teacher to explain it. And then you had the issue where there were teachers who really didn't believe in it yet and didn't have very good things to say about it to the students. So what are kids supposed to think if the teachers are being pretty negative about it or, or just can't really explain it that well? But more so for the top learners who had drove them crazy that they got a three and worked very, very hard and there's somebody who maybe didn't work quite as hard or whatever and still got a three or <laugh> or because of the way our "3-2-1" translates to letter grades, a two is a 66% and it's like, so I can either get a 'D' or an 'A'. There was a lot of that at the beginning, like, why can't I get a B or a C anymore? So I think they've been able to let that go quite a bit because after a few summatives, your grade is gonna average out to being something and it's likely a B or a C. And so then that makes them feel more comfortable again. And I think we've just moved a long way from like a '2' isn't a 'D'. A '2' in this criteria means you're proficient but you're not to mastery yet and this is what it would take to get you to mastery. And so I think if we had maybe understood that ourselves a lot better from the get-go and explained that to kids better, we wouldn't have had such negative reactions to it.

Despite acknowledgement of early student confusion and frustration with the grading reform at SHS, Karls stated that he believed there was an overall level of growing student acceptance with the new Targets-based Grading system. Karls indicated that the transition to Targets-based Grading has been an overall positive experience for him and his students. Likewise, he indicated a general positivity for students as it related to their learning and overall acceptance of the changes:

Seeing kids who are maybe not good at math or hated math, or not good at science or hated science in the past, all of a sudden gain an interest of it when they're in your classroom, they either gain an appreciation for the topic or gain the confidence in the topic. That's a good thing. Our grading, the impact it is having on kids, I think that's also been a good thing. Everything we're doing, it's simplified my grading, it's made my assessments much more streamlined. You know, when I look back at old assessments like chapter tests, you might assess the same target four times in a test. It's streamlined things for me and my students. You make sure your targets are aligned to your pedagogy and then therefore aligned to your assessment. It's made things more efficient. The grading aspect itself...I still don't know if I've tackled that yet.

Karls acknowledged that despite perceiving that students were being positively affected by the implementation of the Targets-based Grading as intended by the district, he was still not comfortable with the execution and perceived a mismatch between grading on a 3-2-1 scale and determining mastery level proficiency for students and subsequent impact on students:

I'm still struggling with it. I'm still like how we're, you know, we write our targets so therefore we know what we're teaching. We work through the gaining of the knowledge, whether it's through discourse or presentation or experiments. But when the grading thing

comes around, I still struggle with how do I write a good assessment, make it a 3-2-1 and what is mastery? Another thing I'm having trouble with, and I don't know if this is from district to district, but we're told that, you know, there's discrepancies on spiraling. I tell you what I mean, there was a time in our district we were told, well, we're not supposed to spiral assessments once they've mastered it. But I've seen through my whole career and not argue with anyone, if I give a kid a question that was presented three weeks ago, are they gonna remember it? If they don't, I don't consider that mastery.

Although Karls was positive and passionate about students, their learning, and his teaching practices throughout the entire conversation, and he believed that the shift in grading practices was positively affecting students' school experience, he frequently expressed doubt regarding his ability as a teacher to validly, accurately utilize the Targets-based Grading system to represent student learning.

Teacher Perceptions of Initial Struggles of Implementation

Both Karls and Clark noted some negative staff perceptions throughout the early professional development and preparation stages as it related to standards and subsequent learning targets creation. Clark spoke specifically to the issue of clarity and the difficulties of the initial Targets-based Learning/Grading process SHS went through during the early stages of implementation:

When we really started to implement, we started with standards, breaking down the standards and that was a lot of PLC (professional learning community) meetings that were frustrating because we would look at the common core standards that we were already familiar with and we were supposed to be using and we're trying to break those down. And it was just hard. We just kept asking for examples like what, how do you

break this standard down? What would it, what do you mean by learning target? There was just a lot of, lack of clarity in the process. And then I remember starting to write targets that they wanted to evaluate with their checklist. And I felt like really proud of the target that I had written and I thought I had done it right and it kind of got ripped to shreds in a PLC meeting and I was very offended and left feeling like, “Hey, I'm trying here!”

Clark further indicated difficulties related to the initial creation of the learning targets in the beginning implementation stages at SHS and suggested that although the requisite task of creating targets was largely successful throughout the building, the fidelity of the targets themselves and lack of administrative oversight.

I think people wrote a whole bunch of targets and put 'em in the folder to check the box that they had to do and then realized that a lot of those were garbage and they are, they're just gonna end up redoing it anyway. So I don't know what a better way would've been <laugh>. Um, they had to get, get people to get them done, but um, there's no way you can check everyone's and support everyone's either. So everyone was like, well I did it. Why didn't you tell me they weren't good six months ago? Well, nobody looked at them.

Karls spoke to his personal struggles implementing the changes of the Targets-based grading system even though he was in support of the change:

I feel every initiative that's come through has been something that you can gain from it. I know maybe earlier in my career I was a little more reluctant to change, but whenever something comes through, I try to give it the benefit of doubt. The hardest part I got out of it...was I reluctant? No, no. I've always tried to play along with the crowd and not be too insubordinate. The challenge I had is when I'm assessing a target, how do I challenge

the higher end kids within an assessment with higher end thinking questions and still put it back into a 3-2-1 basis.

Despite some personal challenges with the grading system, Karls noted the professional satisfaction and gratification he felt when seeing firsthand the impact the new system was having. Both Clark and Karls acknowledged a wide variety of initial struggles among their colleagues in early stages of implementation as it related to targets themselves, teacher attitudes and buy-in, and varying understanding in general of the Targets-based grading initiative.

Teacher Perceptions of Growing Acceptance and Shifting Culture

Despite the initial trepidations and stumbling blocks, Clark indicated a belief that there is a growing acceptance and comfort with the Targets-Based learning/grading approach at SHS partly due to efforts from teachers throughout the entire school district:

I think the middle school teachers are doing a really great job with target based and kids are coming up to us now much more comfortable with it and we're just facing a little less resistance with the whole system and that's making things easier for our high school teachers. Um, so yeah, I think over time it'll only get better.

Karls and Clark both spoke optimistically about the continuing district growth and acceptance of staff of the Targets-based Grading reform. They both believed that their colleagues at SHS are generally becoming more comfortable with and positive about the grading reform.

Teacher Perceptions of Legacy Practices

Clark indicated some legacy teacher grading practices that had widely been used by her colleagues within her building were not valid indicators of learning and that they were part of a culture that students knew and accepted. Clark and Karls both acknowledged a strong past culture of traditional, legacy grading practices at SHS. Clark suggested that extra credit and

additional classwork functioned as a means to manage student classroom behavior and keep students busy during periods of classroom downtime:

I picked up on the extra credit being like a big part of the culture here. Like, yeah, well you can offer them extra credit and the kids who did extra credit were either the kids who already had 'A's, and it didn't matter or they just needed like to get over the hump from a 'B+' to an 'A-' or something like that. They were looking at the grade, they wanted to see the letter that they wanted to see. And so I've like rolled with that and I would, I had like quarterly extra credit packets that they could do. And that was kind of also a classroom management tool for me that was like, if they're done early, then I asked them to get out there extra credit packet and be working on that. And that kind of fizzled out and it wasn't very useful. And then we started having more conversations as a staff about like why extra credit isn't really good evidence of learning and, and using things like there was also a culture here for like grading behavior. Like did you get your syllabus signed? That's 10 points. Did you cover your book? That's 10 points. And I was never super into that, but I know a lot of teachers who like the first couple of grades in the grade book every year for kids were whether they covered their book.

As part of the transition to the Targets-based Grading system, SHS administration addressed the culture of extra credit points and even prohibited traditional points-based extra credit for students prior to the 2020–2021 school year. The change in this practice was communicated to students and families via FAQ document which included the following information (seen below in figure 4.17):

Figure 4.17 *Southeastern High School FAQ Addressing Extra Credit*

➤ Why is extra credit not given?

- All scores are based on the students' summative assessments that correspond to the individual targets. Extra credit in the traditional sense is not connected to mastery of a target, however students are encouraged to improve their grade by reassessing a target.

Clark indicated that she did not believe that many teachers, if any at all, at Southeastern were currently awarding traditional extra credit points to students, nor did she believe that students were disappointed with or negatively affected by this change. She believed that students quickly adapted to the elimination of traditional extra credit points and that it did not have an overall impact on the students' learning experiences due to the overall changes in assessments, reassessments, and grading pedagogy.

Student Perceptions

I had the opportunity to meet with and interview six students at SHS. These upperclassmen had all experienced both the former, traditional grading practices at SHS and for the past two school years, the reformed, Targets-based Grading system. The SHS students were collectively energetic, talkative, and all six of them were participants in clubs, sports, and/or co-curricular activities.

Student Perceptions of Positive Impact of Targets-based Grading

The students at SHS overall spoke favorably about the Targets-based Grading system in relation to its impact on multiple school experiences including classroom learning experiences, time spent connecting with teachers, and homework. One student also recognized the value of having extended time for learning in the event of a school absence:

I just like that it, it helps you like, it makes you feel like you understand what you're doing, and I really do like that you're able to retake the test cuz before that wasn't really an option. Like maybe you just, you missed a day of school or you missed a couple days

and you missed part of the learning. Like you can go back and relearn that and the teacher's there to help you so that you can improve that grade. (SE5)

A common theme among the SHS students was the additional time they were able to spend with teachers under their newer, reformed Targets-based Grading system. One student indicated that although they were initially apprehensive of the new grading system, they eventually grew used to it and grew to like it:

Yeah, I think the new one does cause you get a lot more like one-on-one time with them because like we are this thing called connect time or we can like go to a teacher's class and get help and they'll like really like sit down and explain it to you when you take a test and you'll, they'll like help you through the work you have to do to retake the test and I think that helps a lot. I think there's like benefits and downfalls to both but I think overall, like I didn't really like the new one at first but now that I've gotten like accustomed to it, I really like it. (SE1)

Students recognized the time that teachers initiated, as well as the schedule changes itself that supported this additional teacher attention. Students appreciated less emphasis on graded homework at Southeastern under the new system, and they identified positive changes in reassessment and its impact to allow students time to reach proficiency as it relates specifically to the eligibility components of grades and athletics:

I'd say the benefits are retaking the test. So if you, because I know I like uh, having my grades since I'm an athlete and stuff, you need to have your grades to play in sports. Um, so it's really nice to be able to retake the test and spend time with the teacher and relearning all the stuff so you can get a better grade than the test. (SE2)

In addition to the athletics eligibility factor, another SHS mentioned the aspect of being a student-athlete and being involved in multiple co-curricular activities and the associated time required to be involved in those functions was easier to manage under the new reformed system.

Student Perceptions on Retakes and Reassessment

The Southeastern students indicated that the greatest impact to their overall school experience was in the form of the changes in teacher and building assessment practices. Students indicated that the new culture of assessment and reassessment, or “retakes,” was one that reinforced multiple opportunities for students to show their learning, and subsequently, the new grading system reflected and reinforced that change. One student shared an opinion that they believed the new system reflected actual learning better given the narrative descriptors of proficiency compared to the previous system that provided a grade without any real context to the learning:

I like the retake aspect of it. Like being able to retake the test if maybe you didn't quite understand what you were doing or you just weren't quite ready for the test is super helpful. I think this one would reflect your learning better just because it, it shows that you either understand it, you almost, you almost have it or you just, you're not quite there yet. Whereas the other one it was like, oh, I gotta be like, it doesn't, it doesn't really mean anything to you whereas like this is showing that you either have it or you don't. (SE5)

Southeastern High School’s approach to reassessment opportunities prior to final grades being provided to students is was communicated to all stakeholders prior to the 2020–2021 school year as evident through the reassessment for grading expectations in figure 4.18 below:

Figure 4.18 *Reassessment for Grading Expectations*

- Prior to reassessment, students will need to ensure that new learning has occurred, which places an emphasis on the process of learning.
 - Re-teaching
 - Homework completion
 - Practice

- The reassessment may be one or more of the following:
 - A different assessment
 - A portion of an assessment
 - An opportunity to improve a project
 - A conference with the student

- Students do not have to retake an entire assessment. If he/she struggles on a particular concept or skill, he/she is allowed to retest in that area to replace the score assigned to that particular target.

- The most recent evidence of mastery will be graded and reported, even if the student scores lower on the reassessment. However, students may have multiple opportunities to reassess until they reach mastery.

It was clearly evident throughout the interviews that students at Southeastern consistently praised the SHS Targets-based Grading system's approach to retakes and multiple opportunities to show understanding.

Student Perceptions of Homework and Legacy Practices

Students acknowledged that few legacy practices familiar to them such as extra-credit and being penalized for behavior issues still existed at SHS, and they acknowledged the change in grading system extinguished most of the traditional practices. Students also indicated they felt that the reduction of assigned homework, graded homework and increased in-class practice work was a welcome distancing from a more traditional approach to homework.

I'd say there's the same amount of homework. At the same time there's no homework in some classes now that with the 3-2-1. Which is why with the tests, all the tests basically are 80 percent of your grade. Then homework is like 20 percent. So if you really, if you do good in the test, that's gonna bump your grade up pretty good if you do bad if it's

gonna take a toll on your grade because if you're doing, if you're not doing too well.

(SE2)

Some students indicated that although the new grading system was in place and fewer legacy practices were happening at SHS, students were still motivated by points. Although the current grading system is no longer based on 100 points scales, students were motivated similarly based on 3 point scales.

Student Perceptions of Change in Teacher Practices

Another common theme among the Southeastern students was their recognition of the intentionality behind the teacher practice of breaking assessments into smaller, more frequent tasks which subsequently has led to more manageable and targeted opportunities for retakes and grading proficiency improvements:

I think the new one also does a better job. Because it's just like, it like targets certain things instead of just like giving you like a, like a wide like variety, like do this specific thing and if you do it well, you get a good grade. If you don't, then you can redo it until you do know it. Like the assessments are like smaller, like they're broken up smaller than they were before. Because it's like you're doing one thing instead of like three of four, like for a big test. (SE3)

In addition to recognizing the benefit of teachers breaking assessments into smaller chunks, students also recognized the benefit of the use of learning targets themselves. One student suggested that teachers have stayed true to the learning targets in terms of class learning expectations and assessment content:

I think this one (grading system) is a lot easier to understand because you know what you have to learn cuz like before we learn a lesson they like will give you the target like wrote

out and then you like kind of have an idea of what you're gonna learn and what's gonna be on the test because it and a lot, like in almost every class it stays really true to that target what they gave you before. (SE1)

Students recognized that some teachers underwent a change in instructional practices and instructional delivery as a result of the Targets-based Grading reform. They indicated that in the previous system, students were stuck with their initial test grade without any additional opportunities to show their learning and subsequently improve their grading:

Well I think it's easier to like retake stuff. Like if you don't do good the first time, you can go back and do it again. Because before that some teachers like they don't do a lot of retakes. So like you get that test grade, and you're like stuck with it. But with this you're not. (SE3)

It was apparent that SHS students recognized changes in their building schedule, teacher practices, and overall grading system were intended for their benefit and learning.

Student Preferences with Grading Systems

Student feedback and opinions of preference regarding the two grading systems they experienced throughout their high school careers was varied. Although some sentiment suggested that the Targets-based System is more straightforward in terms of whether students know the learning tasks and their grades reflected such, other comments indicated that students preferred the more comfortable, societally ubiquitous traditional letter grade approach:

I'd say I prefer the previous grading system because like whatever grade you got, you got, so if you studied you would get a good grade. If you didn't, you didn't. And now it's kind of like you don't know what grade you're gonna get anymore. (SE6)

Two students expressed that they liked the newer Targets-based Grading system because they believed that this newer system did a better job with facilitating the understanding of their actual learning:

I feel like that they both have like, their good and bads like this one, it's easy to understand. Like you either know what you're doing, you're like almost there or you just don't have it. Whereas the other one? You know, it was such a wide spectrum of where you could be at. (SE5)

Similarly two Southeastern students highlighted pros and cons within both systems and likewise recognized the benefit of the Targets-based Grading system breaking learning down into perceivably smaller, more manageable chunks; nonetheless, both students expressed an overall general neutrality in terms of their preferences between the two systems:

I would say with ours now, the Targets-based, we communicate with our teachers a lot more cuz we need to be able to retake tests and like they kind of just explain like what grade you would need to get on each test in order to get a three or a two or whatever. Um, it was kind of, it was just a big change. Um, I like it for my sake cuz I like to be able to retake things, but I also have gotten twos on some things that I really should have gotten a higher score on, so I like also don't like it. I am kind of indifferent. (SE6)

I feel like the new system, it's better in a way, but at the same time the old system was more like we got more time to work on notes and stuff like that. There was more stuff in the topic and like we could, we would pack it all together, which is at the same time struggling to um, like if you're on the end of the test at the end of the chapter there's a lot you have to relearn and everything. But with the 3-2-1 grading scale we basically just study one topic, you'd take a test on it, then you move to the next. It's, hmm, it's,

there's a lot of differences and um, similarities to it, but at the same time I feel like the old grading is a little better to me. (SE2)

Students at SHS were cognizant of the overall increased learning that was happening largely as a result of the change in grading systems and assessment philosophy associated with the Targets-based Grading system.

For my actual learning, I'd say I prefer the one we have now with the Target Based grading, because you're supposed to like be able to really understand the material before you take the test. So that's why we do retakes. You have to like do some requirements before you do a retake. So it's like really helping you understand what you're doing.

(SE6)

Student interview data suggested that they understood and generally embraced the four tenets of Targets-based Grading that SHS communicated to all stakeholders at the start of the 2020–2021 school year related to proficiency of learning objectives, less work receives a grade, multiple opportunities for students to learn does not impact full credit, and grades are no longer intertwined with behavioral factors such as participation and work completion (as seen below in Figure 4.19).

Figure 4.19 *Tenets of Targets-based Grading*

TARGET BASED
Grades are based on a 0-3 scale that measures a student's level of proficiency towards clearly articulated learning objectives.
Much less work receives a grade. On work that is for practice and growth, teachers provide feedback only.
Failure to "get it" early on does not impact student grades. If a student demonstrates improved mastery of a concept, then they receive full credit for that understanding.
Grades are NOT used for punitive purpose. The grade should only represent a student's level of proficiency. Therefore, factors like effort, participation, punctual work completion, and time management are not incorporated into a student's grade.

Although there were varying student preferences between the former and current grading system, SHS students expressed a genuine understanding of the Targets-based Grading system and spoke positively of both their former, traditional system and current grading system.

Chapter 5:

Discussion and Conclusions

My research sought to understand students' perceptions of grading reform. The findings of this chapter are organized and presented around the following research questions:

1. How has transitioning to Standards-based Grading reform affected students' levels of perceived interest in academic coursework and school experience?
2. From a student perspective, what “legacy practices” in grading still exist, if any, within a Standards-based Grading system?

The conceptual framework as described in chapter three centers on the concept of students standing at the epicenter of grading practices affected by teacher practices, building expectations and policies, and systemic changes related to grading. I compare the findings from both Northeastern and Southeastern high schools, lay out a comparison of cases, and discuss the overall conclusions of the study. I present the three primary themes that emerged from the two high schools in this research study related to the research questions and conceptual framework. Interviews with students, teachers, and administrators were primary data sources that addressed these questions in triangulation with myriad district documents and artifacts from both sites as well as limited site observations.

Themes of Findings

The coding process helped me strategically place similar codes from the stakeholder interviews into subcategory clusters which ultimately led to my three primary themes. A theme can be an outcome of coding, categorization, or analytic reflection (Saldana, 2016). Data from all stakeholder interviews was used throughout the coding analysis process and theme identification

process. Saldana (2014) indicated the usefulness of using categories during the qualitative research process:

Category construction is our best attempt to cluster the most seemingly alike things into the most seemingly appropriate groups. Categorizing is reorganizing and reordering the vast array of data from a study because it is from these smaller, larger, and meaning-rich units that we can better grasp the particular features of each one and the categories' possible interrelationships with one another. (p. 14)

The primary themes of findings (Comfort, Academic Relevance, and Perceived Fidelity of Implementation) with respective subcategory clusters are as follows:

Figure 5.1 Primary Themes of Findings		
<i>Theme 1: Comfort</i>	<i>Theme 2: Academic Relevance</i>	<i>Theme 3: Perceived Fidelity of Implementation</i>
Preferences	Points; "points game"	Potential
Dissimilarity	Rubrics	Pushback
Familiarity	College Admissions and Relatedness	Unprofessionalism
Fairness	Feedback	Inconsistency
Confusion	Reassessment	Buy-in
Frustration	Homework	Rushed
	Conversion	Understanding
	Learning	Traditional
	Grade Point Average	Communication

The interview data from students, teachers, and administrators touched on many themes related to a continuum of factors associated with grading reform, grading practices, and implementation processes.

Comparison of School Contexts and Demographics

Each of the schools presented various similarities in demographics, geography, and implementation processes of reform. Both schools were singular high schools located in rural communities of fewer than 5,000 people. The schools had relatively similar numbers of students (254 vs 323), proportion of students of color (12.6% vs 10.5%), students of low socioeconomic status (34.6 vs 42.4%), and disability status (17.7% vs 10.2%). In terms of implementation of Standards-based Grading reform, each school considered themselves to be in their third year of implementation although the processes and pilot of implementation looked different at each site. The overall demographics, implementation timelines, and school contexts of both sites were similar in nature.

Structural/Administrative Comparisons

Common variables and topics emerged from across both sites related to the selected grading system, administrative groundwork and initiative rollout, and initial support developed for professionals.

Figure 5.2 *Comparison of Structural/Administrative Variables*

Variable	Northeastern High School	Southeastern High School
Grading System	<ul style="list-style-type: none"> • Targets-based system following Hierck & Larson (2018) format • Narrative, descriptive indicators of proficiency (e.g., Proficient vs Developing) 	<ul style="list-style-type: none"> • Targets-based system following Hierck & Larson (2018) format • Numerical indicators of proficiency (e.g., 3 vs 2)
Administrator	<ul style="list-style-type: none"> • Primary driver in implementation of Targets- 	<ul style="list-style-type: none"> • Primary driver in implementation of Targets-

Variable	Northeastern High School	Southeastern High School
	based Grading	based Grading
Structural/Administrative	<ul style="list-style-type: none"> ● Staff professional development focused on the “why” of grading transition ● Teacher pushback and resistance during implementation 	<ul style="list-style-type: none"> ● Greater degree of student involvement in early stages of implementation (student surveys) ● Staff professional development focused on foundational work of targets creation ● Teacher pushback and resistance during implementation

Both schools were led by strong, passionate administrators who believed in their respective Targets-based Grading reform efforts and the positive impacts it would have on students, and both administrators were the catalysts for the grading reform in their respective schools. The administrators each reflected beliefs in the importance of allowing students multiple opportunities to practice and demonstrate their learning consistent with best practices literature (Fisher et al., 2011) and the use of multiple grading criteria beyond large stakes standardized assessments (Guskey, 2020; Guskey & Bailey, 2019).

As discussed in the literature review, there are many reasons Standards-based grading should replace traditional grading, and an effective grading system should be fair, reliable, valid, clear in purpose, and useful (Munoz & Guskey, 2015; Scriffiny, 2008). The overall approach, expectations, and systems of grading in both high schools were similar in nature as they worked to implement a more effective grading system. In both settings, schools utilized a Targets-based Learning and Grading iteration of Standards-based Grading, preferring to grade students aligned to teacher-created targets which subsequently aligned to state standards, opposed to a more traditional direct grading alignment to the state, or national, standards themselves.

The administrators from NHS and SHS each indicated a variety of challenges throughout the years leading up to the actual implementation including the background work of creating and revising targets, working with teachers, challenging teacher belief systems, and generating buy-in from staff, students, and the greater community. Similar struggles related to implementation were present across the school sites and were articulated by both administrators which included attempts to extinguish grading practices largely regarded as non-conducive to student learning (Brookhart & Guskey, 2019; Cross & Frary, 1999; Kohn, 2013; O'Connor, 2011) related to homework, extra-credit, mean averaging scores and teacher biases. Throughout their implementation processes, administrators from each site indicated at least some involvement of students in the process; for example, Southeastern surveyed students during the early stages of implementation. However, neither site systematically sought to include students' voices as they worked to implement and transition to Targets-based Grading. Yet, the literature suggests that involving students can improve reform efforts since students can provide important insights that are not otherwise available (Conner, 2022; Kushman, 1997; Mitra, 2018). Each administrator indicated there was room for improvement and/or greater future involvement of student voice and student feedback to their grading processes.

Although the sites followed similar planning, professional development, and implementation timelines, there was a noticeable difference in the style of professional development. As suggested in the literature, implementation of Standards-based grading will invariably be met with unanticipated issues and complexities (Kalnin, 2014). This was the case at both sites. At NHS, much of the professional development and staff discussion happened early on, and was continuously revisited throughout the more than 12 months of planning and implementation. These discussions were in relation to the "why" of grading transition, purposes

of grading, pros and cons of grading philosophies, and existing teacher belief systems of the Targets-based Grading philosophy. In fact, multiple staff meetings held throughout the process focused almost solely on those teacher-related beliefs and philosophies.

In comparison, SHS spent significantly less time on staff beliefs and purpose of transition in grading practices and jumped much more quickly into the foundational learning targets creation which is necessitated by the Targets-based Grading system. Guskey (2020) indicated that in regards to grading reform, “commitment to a new approach, confidence that it will work, and trust that positive results will follow almost never occur up front, prior to implementation” (p. 71), which is the general approach that SHS chose as opposed to spending comparable time and energy frontloading the “why” and purpose with staff. Interestingly, teacher perceptions were very similar in both sites despite the vastly different approaches to pre-implementation learning and professional development.

As indicated in the literature review, teacher resistance, pushback, autonomy, and individuality related to grading practices are longstanding factors of secondary education and grading (Brookhart, 1993; Erickson, 2010; Reeves, 2008; Schimmer, 2016; Stiggins et al., 1989; Wormeli, 2006; Zoeckler, 2007). Both sites appeared to have significant pockets of initial teacher pushback, resistance, and persistence of traditional practices. It appeared as if the emphasis on continually addressing the philosophy and rationale for the transition to the Targets-based Grading at NHS had comparatively minimal impact on staff.

Comfort

As a high school student, experiencing the transition from a traditional, legacy practice grading system to a reformed system using a Targets-based Grading iteration of Standards-based Grading affects students’ perceptions of various factors related to their academics and overall

school experience. Comfort was one of the three prominent themes uncovered throughout the study that highlights this impact. This included aspects of grading preferences, familiarity and unfamiliarity, as well as confusion and frustration with grading reform.

Figure 5.3 *Comparison of Variables of Comfort*

Variable	Northeastern High School	Southeastern High School
Familiarity	<ul style="list-style-type: none"> • Students more comfortable and familiar with their former, traditional grading system 	<ul style="list-style-type: none"> • Students more comfortable and familiar with their former, traditional grading system
Confusion and Frustration	<ul style="list-style-type: none"> • Greater degree of reported student confusion with Targets-based Grading system • Descriptive indicators of proficiency significantly more confusing for students • Omission of student voice throughout planning phases 	<ul style="list-style-type: none"> • Students understood 3-2-1 numerical system of proficiency of Targets-based Grading • Omission of student voice throughout planning phases
Preferences	<ul style="list-style-type: none"> • Variety of student preferences • Students expressed greater discomfort due to narrative proficiency indicators • Some student indifference between both systems 	<ul style="list-style-type: none"> • Variety of student preferences • Some student indifference between both systems • Students spoke positively about reassessment opportunities
Teacher Perceptions	<ul style="list-style-type: none"> • Students were confused and frustrated in early stages of implementation • Greater degree of student confusion • Student interest in academic coursework same across both systems 	<ul style="list-style-type: none"> • Students were confused and frustrated in early stages of implementation • Students have become generally positive about Targets-based Grading • Students still converting their 3-2-1 scale to the traditional 100 point scale • Student interest in academic coursework same across both systems

Familiarity, Confusion, and Frustration

Aspects of comfort and discomfort related to switching to a newer, less familiar Targets-based Grading system was present throughout the schools. As discussed in the review of literature, teachers need to abandon their longstanding grading beliefs that they had developed as

students themselves for grading reform to be most effective (Allen, 2005). Some teachers at both sites struggled with the discomfort of a new grading system that did not reflect their traditional beliefs and experiences which contributed to students seeing and feeling that the implementation of grading reform at the respective sites had challenges.

Students also indicated a range of responses such as confusion, frustration and questions about fairness related to the newly implemented grading system. Understandably, students interviewed were nearing the end of their K-12 experience, and much of their time in school was immersed in a legacy system. However, given neither school included students in the planning, pre-implementation, justification for and the “why” related to change, and explanation processes (although SHS involved students to some degree by gauging their concerns and perceptions throughout the early stages of implementation), they did little to build depth of student understanding around the new system. The omission of student voice during the transition process likely further compounded student experiences resulting in confusion. The research reviewed related to student voice is affirming that not only is student voice becoming increasingly popular in education, but high schools with greater student-reported sense of belonging and student voice were correlated with a multitude of positive student outcomes (Conner, 2022; Hall, 2016; Kahne, 2022).

Students at NHS expressed a greater level of discomfort with the grading reform which appeared to stem from their greater degree of difficulty understanding their school’s respective implementation and their perceptions of inconsistent teacher implementation. NHS’s use of its descriptive, narrative indicator of proficiency (e.g., Proficient vs Developing) appeared to be more confusing to their students compared to the simplistic, numerical system (3 vs 2) for SHS students. Similarly, they also expressed a significantly greater degree of confusion with their

respective new Targets-based Grading system than students at Southeastern indicated with theirs. The root of their confusion appeared to stem from their perceptions of inconsistent and dissenting implementation from individual teachers coupled with their perception that the traditional 100 point numbers scale and ABC grading monikers made more sense than the new descriptive target language in their iteration of Targets-based Grading.

Grading Preferences

There was a great deal of variability across student preferences at both Northeastern and Southeastern High Schools related to the traditional grading system they were more comfortable with and used to versus the newer reformed Targets-based System that had been implemented their previous school year. Some students at each site even indicated indifference, as well, not having any strong preferences of one grading system over the other. Interestingly, the majority of students at both sites acknowledged that the Targets-based Grading system likely had a greater impact on their actual learning but that did not always lead to a preference for that system. In some cases, students across the sites indicated that ease, predictability, and familiarity of the 100 point traditional system led to a greater overall preference even while acknowledging the newer, reformed Targets-based System was likely superior for aspects of actual learning. Overall, students did not indicate a significant relationship between their preference of grading system and their interest in their academic coursework as students indicated they were motivated to do well academically regardless of the grading system in use. This is consistent with literature findings that suggest that students tend to conform to whatever grading system to which they are subjected (Cross & Frary, 1999)

In instances where students indicated an acceptance or even preference for the newer Targets-based System, they had no faith nor belief that their post-secondary experience next year

in college would be similar to the new reformed grading system, so they questioned why the district, and why they as students, should bother with the reform if they are just going to revert to a more traditional grading approach next year in college. Similarly, students and teachers alike at both sites suggested that given the fact there is no national, standardized transition to Standards-based Grading reform, the hesitation from stakeholders to embrace the change is understandable and justifiable. This argument from stakeholders illustrates why proponents of Standards-based Grading reform have been urging educators and public education institutions to adopt grading reform since the advent of the Common Core State Standards, and some even prior to that (Guskey, 2009; Hierck and Larson, 2018; Iamarino, 2014; Marzano, 2010; Moss and Brookhart, 2012; O'Connor, 2017; Westerberg, 2016). Nonetheless, there does not appear to be any type of broadly implemented, standardized Statewide or Nationwide reform system on the horizon.

Although the overall preference of grading systems at both sites seemed inconsistent, students at Southeastern particularly spoke positively of their reassessment opportunities and the scheduling structure that supports those opportunities. Students at SHS were generally more positive about their Targets-based Grading experiences compared to student data from NHS.

Teacher Perceptions of Student Confusion and Frustration

Teachers from both Northeastern and Southeastern alike suggested that students, especially in the early stages of implementation, did not understand the Targets-based Grading reform, which led to not only student frustration, but parent and community confusion and frustration as well. Much of this frustration still lingers present time, particularly at NHS. The resulting student negativity and frustration also appears to be in part a result of teacher attitudes and frustration of the grading reform which is not particularly surprising given stakeholders generally do not agree on the overall purpose of grading (Guskey, 2002). Given the degree of

confusion, teachers at SHS indicated that students were generally positive about the grading changes compared to NHS teachers who perceived a great amount of student confusion.

Teachers from both sites indicated that the responsibility of communicating with students was largely left up to individual classroom teachers to explain the Targets-based Grading practices to their students which was a significant factor of student confusion and frustration. At the same time, extensive district documentation from both sites which included handbooks, parent and family communications, FAQs, staff meetings minutes and slides, and district websites appeared to be in stark contrast to teacher perceptions related to causes of student confusion. As discussed in the literature review, a necessary criterion that must be met to ensure Standards-based Grading serves its intended purpose is that (targets) standards should be clearly communicated to all stakeholders—students, parents, teachers—which helps clearly define goals and accurately assess whether students have met said goals (Hany, 2016). It seemed that both districts’ intentions and attempts at communication to all stakeholders were aligned with not only explanatory and procedural information related to the Targets-based Grading, but rationale explaining the purpose behind the transition. NHS’s schedule, timeline, and consistent messaging around the rationale and purpose of grading was extensive; nonetheless, confusion and frustration persisted.

Teacher Perceptions of Student Preferences

As discussed in the literature review, difficulties with grading can be expected when teachers try to merge both a Standards-based Grading approach with a traditional 100 point percentage (Simon et al., 2010). Teachers at both sites indicated initial student dislike for and confusion with the new grading system, and the source of much disdain appeared to be related to an ingrained 100 point scale lens that led to students, and some staff, to attempt to

mathematically compute and translate the 3-point proficiency scale (1, 2, 3 at SHS; Advanced, Proficient, Developing at NHS) into an equivalent 100 point scale resulting in students perceiving their grades were either a 100%, 66% (which also translated in a 'D' grade), or a 33% (which they also translated into an 'F' grade). This conversion process was happening at a much greater degree at SHS compared to NHS based on the fact that SHS utilized a 3-point system that naturally lent to a conversion; contrarily, the descriptive scale at NHS did not. As discussed in the literature review, this mathematical translation is inappropriate and invalid in grading (O'Connor, 2011; O'Connor, 2016) but it was happening among some staff and students at SHS.

Over time, SHS teachers perceived that student preference seemed to be changing given the positive outcomes associated with dedicated changes related to multiple student reassessment opportunities across all classes. This shift, however, appeared to have much to do with the mathematical increases in the reformed grading 3-point grading scale. This attitude suggests students were still motivated to increase their aggregate 3-point score in similar ways they were previously motivated to increase their aggregate 100 point score in the previous system. As suggested in the literature review, instead of incorporating a traditional 100 point grading scale into Standards-based Grading reform, districts should replace the traditional grading altogether (Scriffiny, 2008). It appears that the SHS student mathematical affinity for the 3-point grading actually stems from its relatability to the previous 100 point traditional system. Teachers at both sites suggested that student interest in academic coursework was consistent across both grading systems; likewise, student motivation, albeit driven by different point systems with some students, was consistent across both systems over the past several years of reform.

Academic Relevance

The second prominent theme that emerged was that of academic relevance. Comments, dialogue and data related to the overarching cluster of academic relevance included categories such as homework, rubrics, points and conversion of grades, feedback from teachers, learning, assessment and reassessment, and “points game” (a moniker that represents extra credit, grade changing, points manipulation, and other non-essential to learning factors but nonetheless affect a grade).

Figure 5.4 *Comparison of Variables of Academic Relevance*

Variable	Northeastern High School	Southeastern High School
Impact on Learning	<ul style="list-style-type: none"> • Students perceived Targets-based Grading system (and related changes in practice) was likely impacting learning 	<ul style="list-style-type: none"> • Students perceived Targets-based Grading system (and related changes in practice) was likely impacting learning
Reassessment Practices	<ul style="list-style-type: none"> • Students perceived reassessment opportunities as positive 	<ul style="list-style-type: none"> • Students perceived reassessment opportunities as positive • Structural, schedule change supporting reassessment practices • Additional one-on-one teacher time for students
Impact on Students	<ul style="list-style-type: none"> • Teachers/Administrator believed Targets-based Grading has had positive impact on students 	<ul style="list-style-type: none"> • Teachers/Administrator believed Targets-based Grading has had positive impact on students
“Hodgepodge” Grading	<ul style="list-style-type: none"> • Strong, written expectations prohibiting “Hodgepodge” practices (e.g., extra-credit; student behavior) • Only minimal evidence of “Hodgepodge” grading practices in use 	<ul style="list-style-type: none"> • Teacher perception of strong previous culture of “Hodgepodge” grading, particularly extra-credit

Student Perceptions of Grading Reform Impacting Their Learning

As indicated in the literature, it is important for teachers to allow students opportunities to practice the formative classroom work without a connection to a grade; likewise, multiple

opportunities to demonstrate learning based off of formal, summative assessments should be part of the reformed grading system (Fisher et al., 2011). Perhaps the most profound impact that the transition to a Standards-based Grading reform had on students academically was in regards to their perceptions of actual learning which was highly related to assessment practices. By the second year of implementation, it was apparent at both schools that students recognized the changes in grading system, and subsequent changes in correlational assessment practices, was likely leading to longer-term, improved learning. In fact, students at both sites indicated that they had been driven to get good grades in the previous system regardless of whether actual learning was happening. It seemed apparent that some students at both sites were playing the “points game” in the previous system and were chasing points and grades instead of being motivated by learning. At both sites, however, students indicated that the emphasis on multiple opportunities for assessment within their reformed grading system was a significant factor to this shift in their beliefs about their own learning. Both sites adhered to best practices aligned to summative reassessment.

The literature indicated Standards-based Grading was a means for allowing students an active role in monitoring their own learning and growth (Knack et al., 2012). While this was not evident at NHS, students at SHS recognized that the change in grading system and subsequent, necessary change in instructional practices led to more one-on-one time with teachers due to the reteaching and reinforcement of learning that was associated with the opportunity to reassess. Students were frequently monitoring their learning and grades, and they initiated efforts with teachers for reassessments and subsequent improved grade outcomes. Students were aware of how they were performing as measured against the learning targets. This change was a powerful

outcome of the grading reform at SHS because it generated staff and student buy-in as a result of both stakeholders perceiving it to have a positive impact on learning.

Student Perceptions of Reassessment Practices

Students at both sites perceived that opportunities for reassessment was a positive practice. At SHS especially, the culture of reassessment was an expectation for all staff as they changed their overall support structure by modifying their building instructional schedule to include a period dedicated for reassessing students. This structural schedule change that supported the Targets-based Grading reform had a profoundly powerful impact on the building culture, teachers, instructional practices, students, and student learning. Although the review of literature did not address any formal building schedule or structural designs related specifically to reassessment, grading purpose linked to multiple summative assessment opportunities and evidence of learning outcomes has wide support in the grading reform literature (Guskey, 2020; Guskey & Bailey, 2009; Guskey & Link, 2019; Guskey, 2006).

Students at Southeastern articulated that within their Targets-based Grading system the emphasis for actually learning the material was the impetus behind the emphasis on reassessments and multiple opportunities to demonstrate learning. Students also recognized secondary benefits from the emphasis on reassessments including that student athletes would have multiple opportunities to improve their proficiency and hence academic eligibility for participation. Furthermore, students at both sites recognized that multiple opportunities for reassessment and demonstration of learning led to grading outcomes that they felt more in control of compared to the previous, traditional system that they perceived was a “single assessment, teacher moves on” type of system. However, not all students saw the ability for unlimited retakes as a positive. One student expressed a belief that having unlimited

opportunities to reassess could mean that students would blow off the first couple opportunities which likely did not foster good initial preparation or study habits.

Teacher/Administrator Perceptions of Grading Impact on Students

Teachers and administrators from both sites also expressed belief that the Standards-based Grading reform had a positive impact on students' academic and school experiences. Examples were cited from initial implementation of students not doing homework or other formative work because they did not see a direct line connecting their ungraded homework to their course grades. As O'Connor (2011) suggested "when homework assigned as practice is scored and included in grades, what becomes most important to students is that it be done because it 'counts', not because of any learning that might occur" (p. 110); through the Targets-based Grading reform, students, and staff, were experiencing that transition from scored to unscored homework. Over time, however, students began to realize that although homework was ungraded, per se, it was a necessary component of practice and learning that indeed ended up impacting their final grades. At both sites this led to teachers shifting their practices to better entice and motivate students to complete homework and help students see the relevancy of homework as it related to their learning and final grades. The adults recognized that there was a much needed transition from the "points game" culture of students being motivated by points to a culture of students being motivated by learning. This transition started to take place the second year of Standards-based Grading reform at both sites. As discussed in the literature review, the intrinsic motivation for learning is far more valuable than extrinsic rewards and/or punishments (Wormeli, 2018). It appears both sites are gaining momentum and success with helping students find intrinsic value in their learning. Both sites have successfully minimized the grading impact on homework, a practice that has wide support in the literature identified as a means of reducing

non-essential, traditional factors of calculated grading for students (Kohn, 2006; O'Connor, 2011; Schimmer, 2016; Reeves, 2017).

Impact of “Hodgepodge” Grading Practices on Students

Another shift that happened through the grading reform process that impacted students, particularly at SHS, was in regards to teachers awarding extra-credit, a notorious “hodgepodge” traditional grading practice related to grade inflation and invalid, unreliable grade reporting (Cross & Frary, 1999; Erickson, 2011). Students at both sites acknowledged teacher inconsistencies in grading and legacy practices such as extra-credit, aspects related to “hodgepodge” grading; however, students did not necessarily see a problem with extra-credit opportunities which is not surprising given that the literature suggests that students may not understand negative aspects of “hodgepodge” grading (Cross & Frary, 1999). The longstanding culture of extra-credit being provided as a means to bump student grades to the next threshold (e.g., B+ to an A-) was specifically addressed through policy at SHS. District documentation explicitly indicated that all grading scores were to be based on the students’ summative assessments and that extra-credit in the traditional sense was prohibited. Similarly to extra-credit, the culture of grading student behavior saw some significant shifts as well. Grading practices convoluted with student behavior are not in alignment with purposefulness of student performance, growth, or achievement (Deddeh, 2010; O'Connor, 2011; Randall & Engelhard, 2010; Wormeli, 2006). One example cited by a SHS teacher was her colleagues’ practice of awarding points for compliance behavior such as getting a syllabus signed by a parent or covering a textbook with a textbook cover. NHS had explicit, written expectations for staff that behavior must be separated from academic grades. Although NHS appeared to have more teacher resistance and autonomy, there was little evidence of “hodgepodge” grading. This difference is

likely as a result of the strong written expectations to staff that was also communicated to families.

Perceived Fidelity of Implementation

The third primary theme identified through this study is Perceived Fidelity of Implementation, which includes all variables pertinent to the implementation of the grading reform process and the transition from the previous process. This included staff and student pushback, staff unprofessionalism, inconsistency of implementation, lack of buy-in, timing and perceptions of a rushed implementation, and understanding/clarity of grading systems. It was apparent through the interviews and documents review process that implementation and perceptions of implementation fidelity were wide-ranging.

Figure 5.5 *Comparison of Variables of Perceived Fidelity of Implementation*

Variable	Northeastern High School	Southeastern High School
Implementation Process	<ul style="list-style-type: none"> • Students perceived initial Targets-based Grading implementation rollout was problematic • Students indicated inconsistencies between teachers; operational definitions of “advanced” inconsistent across teachers • Teachers and Administrator indicated inconsistency of implementation 	<ul style="list-style-type: none"> • Students perceived initial Targets-based Grading implementation rollout was problematic • Students indicated inconsistencies between teachers • Teachers and Administrator indicated inconsistency of implementation
Accountability	<ul style="list-style-type: none"> • Only one administrator; limited admin time for following up on execution/compliance of grading practices • More prominent issues of implementation present day 	<ul style="list-style-type: none"> • Only one administrator; limited admin time for following up on execution/compliance of grading practices • Level of staff pushback shifted positively due to some teachers leaving
Teacher Autonomy	<ul style="list-style-type: none"> • Students perceived teacher autonomy and inconsistency • Admin noted some teachers still using secret grade books behind the scenes 	<ul style="list-style-type: none"> • Students perceived teacher autonomy and inconsistency

Variable	Northeastern High School	Southeastern High School
Perceptions of Change	<ul style="list-style-type: none"> ● Students perceived individual teacher instructional practices have mostly remained unchanged 	<ul style="list-style-type: none"> ● Students perceived individual teacher instructional practices have mostly remained unchanged ● Students believe previous system more transparent with grades and had more data ● Students perceived Targets-based Grading 3-point system has hidden mechanisms for grading ● Students recognize significant shift in reassessment practices and schedule as a result of new grading system

Student Perceptions of Implementation

The implementation of the Standards-based Grading process at both Northeastern and Southeastern High Schools continues to impact students' perceptions of their school experience. Students at both sites recognized that the implementation rollout was problematic based on their perceptions of staff and student confusion, frustration, and lack of teacher consistency. To date, students expressed they continue to experience inconsistencies between teachers. One Northeastern student indicated specifically that they perceived teachers' application of the district performance grading standards were inconsistent with one another. For example, the operational classroom definition of "advanced" meant one thing to one teacher and something completely different to another. This resulted in students, particularly those who are motivated to achieve perfect or near-perfect grades, feeling frustrated.

Teacher/Administrator Perceptions of Fidelity of Implementation

Teachers and administrators at both sites acknowledged inconsistency of implementation and a lack of fidelity particularly in the early stages of Targets-based grading implementation.

Teacher data also suggested that the lack of accountability is a factor in the implementation and

its inconsistencies. The lack of accountability comes from a small number of administrators (only one at each site) who are pulled in multiple directions having little to no time to commit to follow up conversations with staff regarding continual improvement, execution and compliance of intended grading practices. As discussed in the literature review, the model of professional development needs to change with a greater emphasis on individual teacher feedback and coaching of teachers while they implement reform (McMunn et al., 2003). Given the limited administrative time dedicated to monitoring continual execution of Targets-based Grading practices and fidelity at both sites, individual teacher feedback and coaching would likely be impactful. Teachers acknowledged that the sole building administrator cannot effectively be responsible for monitoring continual implementation.

Individual teacher beliefs, practices, and autonomy is a powerful variable as it relates to the fidelity of implementation. Although there was a previously strong history of traditional, legacy grading practices at SHS, the level of staff pushback shifted a bit as a result of some resistant teachers leaving the district prior to the actual implementation in combination with some staff embracing the change. In comparison, NHS issues related to fidelity of implementation seem to be more prominent in the present day than at SHS.

Student Perceptions of Teacher Autonomy

The literature review states that teacher grading practices at the secondary level tend to be subjective, unpredictable, and full of individual autonomy (Duncan & Noonan, 2007; Erickson, 2010; Schimmer, 2016; Wormeli, 2006). Perhaps unsurprisingly, students at both sites were able to point out inconsistencies between their classroom teachers' autonomous practices, and they recognized that some practices were inconsistent with and unaligned with the intended practices of the school's new reformed grading system. This is also consistent with the literature that

suggests teacher autonomy in grading is shown through their individual grading practices, and those practices are taken seriously, personally, and emotionally (Schimmer, 2016). Students spoke of specific teachers who were discreetly still utilizing their traditional 100 point grading system behind the scenes which is also consistent with literature review findings that grading practices tend to be more based on tradition than evidence of actual effectiveness (Guskey, 2020). At NHS, administration noted some teachers were still using “secret little gradebooks” behind the scenes despite the implementation of the Targets-based Grading. Predictably, students and parents discovered this which then led to frustration from multiple stakeholders and required administrator intervention in the form of parent phone calls and explanations at Northeastern. As Reeves (2017) suggested, despite some districts’ best efforts to reform grading, it can be met with resistance and backlash from stakeholders including and especially parents. Comparatively, students at SHS reported fewer teacher inconsistencies but they existed nonetheless.

Student Perceptions of Actual Change

The literature findings suggested older high school students become more cynical of grading systems and find grades to be less meaningful (Evans & Engleberg, 1988). Students at both sites expressed some cynicism as they felt that, despite the change in grading systems on the surface, little had really actually changed from their perspective in terms of many individual teacher instructional practices. In most cases, students believed that teachers largely taught and graded the same but did so within the confines of the new systems. However, at SHS, students recognized intentional changes in shorter chunked assessments and subsequent grading opportunities in smaller, more frequent occurrences. Some SHS students perceived that the older system was more transparent with more grades and more data and more scores while the newer,

3-point systems had hidden mechanisms for determining proficiencies and were less consistent than the traditional system.

In terms of systems change, students at NHS suggested that there was less actual impact to their school experience with the Targets-based Grading reform likely as a result of many instructional and teacher practices remaining the same as they had been in the previous system. Conversely, SHS students perceived that the overall emphasized building shift in reassessment practices was a significant synergistic factor with the grading change that led to a different overall experience. Southeastern students suggested that the change to the daily schedule which resulted in the inclusion of a specific period of the day dedicated to reassessment was a significant factor.

Conclusion

How has transitioning to Standards-based Grading reform affected students' levels of perceived interest in academic coursework and school experience?

At the high schools studied in this multisite case comparative study, the transition to a Standards-based Grading (Targets-based Grading iteration in both schools) system affected students' perceived experiences in a variety of ways ranging from confusion to perceptions of positive impact on learning to beliefs about individual grading practices and autonomy. General concluding outcomes related to students' levels of perceived interest in academic coursework and school experience are as follows:

1. Students adapted to and accepted change in the grading system even if they may prefer the more traditional, more familiar previous system.
2. Students experienced confusion and frustration with the transition to the Standards-based (Targets-based) Grading systems. Over time, as students had

more exposure and experience with the reformed grading systems, they started to embrace some of the changes. Teacher perceptions of student confusion and frustration aligned with student perceptions.

3. Students who experienced the supportive structural building change (reassessment period added to Southeastern High School schedule) indicated a more profound change and impact to their school experience as a result of the Targets-based Grading reform.
4. Student interest in coursework remained constant; likewise, students were motivated to get good grades in both systems. The emphasis on reassessment in the reformed Targets-based system was positive for students and they saw the link between reassessment, improved learning, and potentially improved grades.
5. Students perceived that the Standards-based (Targets-based) Grading systems had a positive impact on their learning even though their overall preferences were varied and some preferred the more traditional system. Teacher and administrator perceptions of positive impact on student learning were consistent with student perceptions.

Students were insightful and introspective when talking about their perceptions and their lived experiences related to their respective school's grading reform; however, both schools missed an opportunity to involve students earlier in the planning process prior to implementation. As Guskey (2020) indicated, "when the [district] considers and values students' perspectives in planning grading and reporting reforms, not only will students better understand the reasons for change, but they will be far more likely to support the implementation" (p. 49). Although student support and buy-in of the Targets-based Grading has improved over the two years of full

implementation, both schools would be remiss to not continually, consistently include students feedback as they continue to evolve and improve their grading systems.

From a student perspective, what “legacy practices” in grading still exist, if any, within a Standards-based Grading system?

At the high schools studied in this multisite case comparative study, some “legacy practices” in a Standards-based Grading (Targets-based Grading iteration in both schools) system still exist, and students are acutely aware of their existence. General concluding outcomes are as follows:

1. Students perceived legacy practices still exist primarily in the form of some individual teacher grading practices (e.g., using 100 point scale). Teacher and administrator perceptions aligned with student perceptions.
2. Students perceived that although some legacy practices still exist, the Standards-based (Targets-based) grading system does not support these practices.
3. Students perceived that individual teacher autonomy is a significant factor in teaching and grading practices. Teacher and administrator perceptions aligned with student perceptions related to individual teacher autonomy and perceived implementation of grading practices.

Individual teacher autonomy is a significant factor in implementation of grading reform. As Townsley (2013) suggested, even when teachers are part of the reform process, their perceptions and personal belief systems are still a factor and may impede implementation. NHS took a less direct approach to minimizing legacy practices and utilized professional development time and resources designed to challenge teacher beliefs and continually bring back focus on the rationale for implementation of new practices. At both sites, teacher autonomy with grading practices is

the likely culprit for the last remaining legacy practices. Teachers and administrators acknowledged that the Targets-based Grading and subsequent classroom practices does not support or condone the legacy practices. However, behind closed classroom doors, some autonomous legacy practices such as traditional points and grades in teacher gradebooks still exist in pockets at both Northeastern and Southeastern High Schools. It is important to note that although some legacy practices still exist at both sites, the transition to the Targets-based Grading systems has significantly diminished many traditional practices that had been in place throughout both high schools.

Revisiting the Conceptual Framework

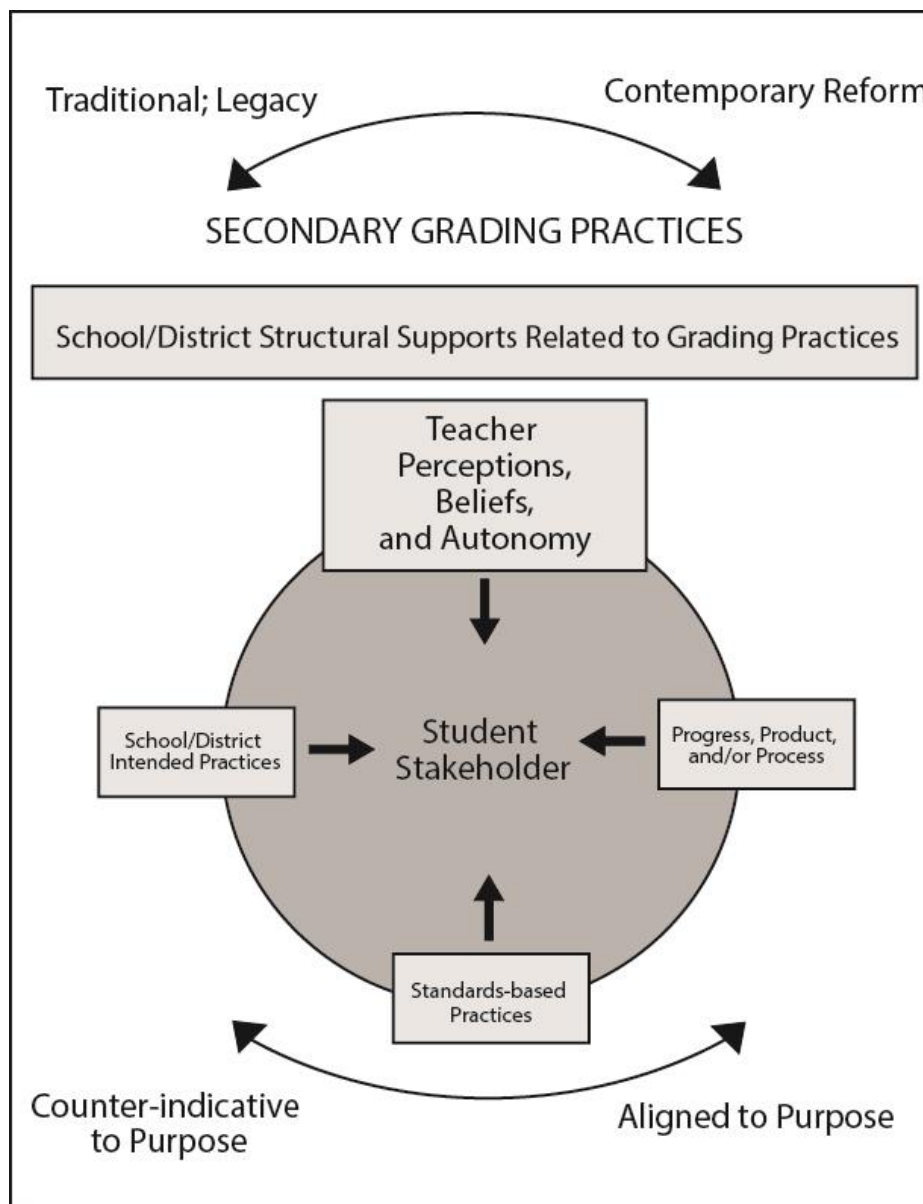
The original conceptual framework presented in Chapter 3 posited that the student is at the epicenter of secondary grading which represents the fact that much of what is done in education does and should center around the student and student learning. This holds true in my revised framework below. On the perimeter of the outer circle of the original framework, five boxes represent major grading practice influences to the student stakeholder: teacher perception of student performance; the progress/product/process components that makeup a grade; Standards-based reform and practices; teacher autonomy, specifically the belief system individual teachers bring to grading; and the school or district's formal, ostensible grading policies and practices. The arrows pointing toward the student stakeholder reinforce the idea of the student being the epicenter of the five variables.

In the revised framework, I have three original boxes (Standards-based grading reform; School/District intended practices; and progress/product/process components). However, teacher autonomy and teacher perceptions have been combined into a much larger 4th box that represents all components of teacher beliefs, autonomy, and perceptions related to grading. The

size and position of this enlarged box rests at the top of the student circle to intentionally denote the role these individualistic teacher attributes play in grading practices at the secondary level based on my findings.

Even greater in size above the entire circle is the box titled “School/District Structural Supports Related to Grading Practices.” This new addition to the conceptual framework reflects one of my primary study findings as exemplified by the changes Southeastern High School made to their daily building and student schedules for the purpose of retakes and reassessments. This change had a significant impact on staff and student buy-in and the overall efficacy of the grading system reform.

Figure 5.6 *Revised Conceptual Framework of Secondary Grading Practices on the Continuum of Change*



Limitations of the Study

This study and conceptual framework centered on student experience and student perceptions. Consequently, the results of this study are thereby limited since it did not focus on another primary stakeholder population in the world of K-12 grading: families/parents. Student perception and student voice are critical components to grading reform, but parent/family voice is arguably just as important. We know parents appreciate and prefer additional communication

from schools (McMuun et al., 2003) which may further indicate that their voice is significant to the implementation of systemic grading changes for their child. Likewise, the school community itself has ancillary staff members beyond just teachers and administrators. Therapists, educational assistants, administrative assistants, health clerks, and other professionals also may have equally important perceptions of grading, grading practices, and grading reform but were not within the scope of this particular study.

The methodological, qualitative approach to this study led to small sample sizes of participants of students, teachers, and administrators. Because of the small sample size of the interviewees, there was a lesser scope of data and information garnered from participants. The sample size scope for students was particularly small given the focus on interviewing students who self-identified as college-bound and who were already academically achieving at a successful threshold as determined by grade point average. Students who had lesser baseline interest in academics or whose paths may be more geared in directions other than continued post-secondary learning may have had significantly varying perceptions of grading reform practices at both sites.

This comparative site study also aimed to answer the previously detailed research questions via multiple data sources gathered in relatively small high schools. These data may not necessarily be representative of larger, more diverse high schools in public education. Similarly, the demographics of student, teacher, and administrator interview participants may not be generalizable to larger, urban communities and districts. The vast majority of all stakeholders at both sites were of White ethnicity and resided in smaller, rural communities.

The student criteria for participation in this study was narrow in the sense that it required them to be upperclassmen who experienced both a traditional approach to grading as well as a

recent transition of Standards-based Grading reform. This meant students were reporting on their experiences of a relatively newly implemented grading system that most likely is still working out lots of bugs and fixes.

Implications for Practice

The impact of student voice in public education can have profound effects on a variety of student outcomes and achievement (Conner, 2022; Kahne, 2022). Eliciting student voice throughout all stages of grading reform—pre-planning, explanation, justification, communication, pilot implementation, full implementation—would be a powerful lever not only in terms of implementation ease, but with stakeholder buy-in as well. Per earlier indication in the literature review, student voice research is a rapidly growing topic of scholarship (Gonzalez, et. al., 2016). Hopefully, student voice research and practice in the area of grading becomes a topic of growth as well. As indicated in the conceptual framework of this study, students are often the passive recipients of educational policies and practices inflicted upon them. This study affirmed that high school students have a wealth of insight, understanding, and feedback from which practitioners would certainly benefit if students were allowed a seat at the grading reform table. Furthermore, student insights may serve to provide valuable stakeholder insights that help inform and steer initiatives such as transitioning away from legacy practices to that of more learning focused Targets-based Grading.

As a practicing school administrator, another part of this study that resonated with me is the need to slow down implementation of change for the benefit of all stakeholders. It was apparent that the implementation of the Targets-based Grading systems at both respective sites would have greatly benefited from a slower, more deliberate implementation process that placed specific emphasis not only on the intended message to all stakeholders, including students, but

ensuring enough time was provided that the message was received and understood and even largely embraced. Stakeholder buy-in of grading reformation appears to be a key component to sustainable change in education. Similarly, helping all stakeholders successfully understand the “why” of grading reform is also paramount to successful implementation.

It was also apparent to me that there were educators at both sites that were not on board with the changes in practice and some of that reluctance was an unwillingness to understand the newer grading system. I believe that any high school investigating grading reform changes would likely benefit from a deep analysis of teacher belief systems related to grading practices. We know that teachers take their grading practices personally, emotionally, and very seriously (Schimmer, 2016). Given the level of perceived staff pushback in this study, I believe districts would be remiss to not spend a significant time helping staff deeply assess and challenge their own beliefs and biases related to grading prior to taking steps towards actual implementation. School improvement is ultimately an administrative responsibility but eliciting teacher buy-in is a major variable to its effectiveness. Grading reform as a means of school improvement is no exception.

We also know that the range and function of grading at the secondary level is vast and wide (Brookhart, 1993; Guskey, 2006; Reeves, 2011). However, the paradigm shift towards Standards-based Grading reform will continue to grow and will continue to be supported by practitioners, authors, and scholars (Guskey, 2009; Hierck and Larson, 2018; Iamarino, 2014; Marzano, 2010; Moss and Brookhart, 2012; O’Connor, 2017; Westerberg, 2016). School districts contemplating grading reform need to consider a variety of implementation factors as they move towards a Standards-based approach. I hope that student perception and voice is consistently one of those factors.

Implications for Further Research

Throughout this study, students, teachers, and administrators shared perceptions of a multitude of factors related to grading at the secondary level as it relates to the reform process of moving from a traditional system to a Standards-based Grading system. Furthermore, stakeholders shared perceptions of how, and to what degree, the grading reform overall impacted the student experience. But one important question beyond the scope of this study that remains largely unanswered is as follows: beyond student perceptions of grading reform on their learning, to what degree has the grading reform impacted actual, measurable student learning? As indicated in the literature review, the vast majority of published literature related to grading lies in the realm of theoretical practice (Brookhart et al, 2016; Olsen & Buchanan, 2019; Townsley & Buckmiller, 2020). Townsley and Buckmiller (2020) stated that “The literature consists of a combination of essays about the problem of classroom grading, surveys of teacher perceptions of their grading practices, and some empirical studies examining the effectiveness of various grading practices” (p. 6). Beyond the continual need to solicit student and other stakeholder voice in grading, more research needs to be done regarding the impact of grading on actual student learning outcomes and data.

Significance

This study contributed to our understanding of student perceptions of grading reform. It specifically adds to the body of research related to high schools moving from a traditional grading system to a Targets-based iteration of Standards-based Grading. This study examined the journey of two high schools on the path of Targets-based Grading implementation and continued improvement of grading practices on behalf of the students they serve.

As a Director of Curriculum & Instruction and central office administrator, this research has helped me personally and professionally. I hope it will help guide my future work with colleagues in my school district as we move closer to implementing an iteration of Standards-based grading. I also hope this study will help inspire school administrators to seek out student perception and input in all stages of planning and implementation of any form of school improvement.

References

- Adrian, C. A. (2012) *Implementing standards-based grading: Elementary teachers' beliefs, practices and concerns*. (Doctoral dissertation). Available from ProQuest Dissertations and Theses database.
- Allen, J. (2005). Grades as Valid Measures of Academic Achievement of Classroom Learning. *The Clearing House*, 78(5), 218-223.
- Anderson, C. (1986). Instructional Leadership Behaviors in High School Principals, Assistant Principals, and Department Chairpersons: A Comparison. *The High School Journal*, 70(2), 115-123. Baron (1999)
- Baron, P. A. (2000). *Consequential Validity for High School Grades: What is the Meaning of Grades for Senders and Receivers?* New Orleans: Paper presented at the Annual Meeting of the American Educational Research Association.
- Black, P & William, D. (2010) Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 92(1), 81-90.
- Bonner, S. M., & Chen, P. P. (2009). Teacher candidates' perceptions about grading and constructivist teaching, *Educational Assessment*, 14(2), 57-77.
- Brimi (2011) Brimi, Hunter M. (2011) "Reliability of Grading High School Work in English," *Practical Assessment, Research, and Evaluation*: Vol. 16 , Article 17
- Brennan et al (2001) Robert Brennan, Jimmy Kim, Melodie Wenz-Gross, Gary Siperstein; The Relative Equitability of High-Stakes Testing versus Teacher-Assigned Grades: An

- Analysis of the Massachusetts Comprehensive Assessment System (MCAS). *Harvard Educational Review* 1 July 2001; 71 (2): 173–217
- Brookhart, S.M. (1993), Teachers' Grading Practices: Meaning and Values. *Journal of Educational Measurement*, 30: 123-142.
- Brookhart, S. (2004). *Grading*. Upper Saddle River, New Jersey: Pearson Education.
- Brookhart, S. (2011). Starting the conversation about grading. *Educational Leadership*, 69(3), 10-14.
- Brookhart, S. M., Guskey, T. R., Bowers, A. J., McMillan, J. H., Smith, J. K., Smith, L. F., Stevens, M. T., & Welsh, M. E. (2016). A Century of Grading Research: Meaning and Value in the Most Common Educational Measure. *Review of Educational Research*, 86(4), 803–848.
- Carey, T., & Carifio, J. (2012). The Minimum Grading Controversy: Results of a Quantitative Study of Seven Years of Grading Data From an Urban High School. *Educational Researcher*, 41(6), 201-208.
- Chappuis, J. (2012). “How am I doing?” *Educational Leadership*, 70(1), 36-41.
- Cizek, G., Fitzgerald S, & Rachor, R. (1996). Teachers’ assessment practices: preparation, isolation, and the kitchen sink. *Educational Assessment*, 3(2), 159-179.
- Conner, J., Posner, M., & Nsoawa, B. (2022). The relationship between student voice and student engagement in Urban High Schools. *Urban Rev* 54, 755-774.

- Cox, K. B. (2011). Putting classroom grading on the table, a reform in progress. *American Secondary Education*, 40(1), 67-87.
- Creswell, J. W. (2013). *Qualitative Inquiry & Research Design: Choosing among Five Approaches (3rd ed.)*. Thousand Oaks, CA: SAGE.
- Crocker J., Quinn D., Karpinski, A., Chase, S. (2003). When grades determine self-worth: Consequences of contingent self-worth for male and female engineering and psychology majors. *Journal of Personality and Social Psychology*, 85, p 507.
- Crooks, A.D. (1933). Marks and marking systems: A digest. *Journal of Educational Research*, 27(4), 259-72.
- Cross, L. & Frary, R. (1999). Hodgepodge grading: endorsed by students and teachers alike. *Applied Measurement in Education*, 12(1), 53-72.
- Cullen, F. T., Cullen, J. B., Hayhow, V. L., and Plouffe, J. T. (1975). The effects of the use of grades as an incentive. *The Journal of Educational Research*, 68, 27-279.
- Cureton, L. W. (1971). The history of grading practices. *NCME Measurement in Education*, 2(4), 1-8.
- Darling-Hammond, L., Hyster, M. E., & Gardner, M. (2017). Effective teacher professional development. *Learning Policy Institute*.
- Denton, J. (1989). Selecting an Appropriate Grading System. *The Clearing House*, 63(3), 107-110.

- De Zouche, D. (1945). "The wound *is* mortal": Marks, honors, unsound activities. *The Clearing House*, 19(6), 339-44.
- Docan, T. N. (2006). Positive and negative incentives in the classroom: An analysis of grading systems and student motivation. *Journal of Scholarship of Teaching and Learning*, 6(2), 21-40.
- Duncan, R. C., & Noonan, B. (2007). Factors affecting teachers' grading and assessment practices. *Alberta Journal of Educational Research*, 53, 1-21.
- Durm, M. W. (1993). An A is not an A is not an A: A history of grading. In *The Educational Forum* (Vol. 57, pp. 294–297). Taylor & Francis.
- Edgar, L., Johnson, D., Graham, D., & Bruce, L. (2014). Student and faculty perceptions of plus/minus grading and its effect on course grade point averages. *College Student Journal*, 48(1), 184-197.
- Erickson, J. (2010). Grading practices: The third rail. *Educational Leadership*, 10(7), 22-24.
- Evans, E. D., & Engelberg, R. A. (1988). Student perceptions of school grading. *Journal of Research & Development in Education*, 21(2), 45–54.
- Fernández-Alonso, R., Suárez-Álvarez, J., & Muñoz, J. (2015). Adolescents' homework performance in mathematics and science: Personal factors and teaching practices. *Journal of Educational Psychology*, 107(4), 1075–1085.
- Finkelstein, I. E. (1913) *The Marking System in Theory and Practice*. Baltimore, MD: Warwick and York.

- Fisher, D., Frey, N., Pumpian, I. (2011). No penalties for practice. *Educational Leadership*, 69(3), 46-51.
- Flyvberg, B. (2011). Case study. In N. K. Denzin & Y. S. Lincoln (Eds.) *The SAGE handbook of qualitative research (4th Ed., pp. 301-316)*. Thousand Oaks, CA: Sage.
- Frary, R., Cross, L., & Weber, L. (1993). Testing and grading practices and opinions of secondary teachers of academic subjects: implications for instruction in measurement. *Educational Measurement: Issues and Practice*, 12(3), 23-30.
- Friedman, S. J., & Frisbie, D. A. (1995). The influence of report cards on the validity of grades reported to parents. *Educational and Psychological Measurement*, 5-26.
- Gonzalez, T. E., Hernandez-Saca, D. I., & Artiles, A. J. (2017). In search of voice: Theory and methods in K-12 student voice research in the US, 1990–2010. *Educational Review*, 69(4), 451-473.
- Gordon, M.E., & Fay, C.H. (2010). The Effects of Grading and Teaching Practices on Students' Perceptions of Grading Fairness. *College Teaching*, 58, 93 - 98.
- Gunter, H. (2012). The leadership of schools. In *Leadership and the reform of education* (pp. 17-36). Policy Press at the University of Bristol.
- Guskey, T. R., & Bailey, J. M. (2009) *Developing standards-based report cards*. Thousand Oaks, CA: Corwin Press.
- Guskey, T. R., & Brookhart, S. M. (2019). *What we know about grading: What works, what doesn't, and what's next*. ASCD.

- Guskey, T. (2006). Making High School Grades Meaningful. *The Phi Delta Kappan*, 87(9), 670-675.
- Guskey, T. R. (2013). The case against percentage grades. *Educational Leadership*, 71(1), 68-72.
- Guskey, T. R. (2009). *Bound by tradition: Teachers' views of crucial grading and reporting issues*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.
- Guskey, T. R., Swan, G. M., & Lee, A. J. (2011). Grades that mean something. *The Phi Delta Kappan*, 93(2), 52-57.
- Guskey, T. R. (2014). Class rank weighs down true learning. *Phi Delta Kappan*, 95(6), 15-19.
- Guskey, T. R., & Link, L. J. (2019). *The forgotten element of instructional leadership: Grading*. *Educational Leadership*, 76(6).
- Guskey, T. R. (2020). Breaking up the grade. *Educational Leadership*, 9, 40-46.
- Guskey, T. R. (2020). *Get Set, Go! Creating Successful Grading and Reporting Systems*. Bloomington, IN: Solution Tree Press.
- Valerie Hall (2017). A tale of two narratives: student voice—what lies before us? *Oxford Review of Education*, 43:2, 180-193.
- Hany, K., Proctor, M. Wollenweber, J., & Al-Bataineh, A. (2016). Teacher perceptions of standards-based grading: Implications and effectiveness. *Journal of Teaching and Education*, 5(1), 749-764.

- Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. New York: Routledge.
- Hierck, T. & Larson, G. (2018). *Grading for impact: Raising student achievement through a target-based assessment and learning system*. Thousand Oaks, CA: Corwin Press.
- Hoover, E. (2012). High-school class rank, a slippery metric, loses its appeal for colleges. *The Chronicle of Higher Education*, 56(15), A1.
- Iamarino, D. L. (2014). The benefits of standards-based grading: A critical evaluation of modern grading practices. *Current Issues in Education*, 17(2), 1-10.
- Johnson, D. W., Skon, L., & Johnson, R. (1980). Effects of Cooperative, Competitive, and Individualistic Conditions on Children's Problem-Solving Performance. *American Educational Research Journal*, 17(1), 83-93
- Johnson, R. T., Johnson, D. W., and Tauer, M. (1979). The effects of cooperative, competitive, and individualistic goal structures on students' attitudes and achievement. *Journal of Psychology*, 102, 191-198.
- Kahne, J., Bowyer, B., Marshall, J., & Hodgins, E. (2022). Is responsiveness to student voice related to student outcomes? *American Journal of Education*. 128(3), 389-415.
- Kalnin, J. S. (2014). Proficiency-based grading: Can we practice what they preach? *AILACTE Journal*, (9), 19-36.

- Kamberelis, G., Dimitriadis, G., & Welker, A. (2018). Focus group research and/in figured worlds. In N. K. Denzin & Y. S. Lincoln (Eds.), *The SAGE handbook of qualitative research* (5th ed., pp. 692-716). Thousand Oaks, CA: Sage Publications.
- Knaack, S., Kreuz, A., & Zawlocki, E. (2012). Using Standards-Based Grading to Address Students' Strengths and Weaknesses. *Online Submission*.
- Kohn, A. (2006). Speaking my mind: The trouble with rubrics. *English Journal*, 95 (4).
- Kohn, A. (1999). *Punished by Rewards:: The Trouble with Gold Stars, Incentive Plans, A's, Praise, and Other Bribes*. Houghton Mifflin Harcourt.
- Kohn, A. (2011). *The case against grades*. *Educational Leadership*, 11(11).
- Krueger, R. A., & Casey, M. A. (2008). *Focus groups: A practical guide for applied research* (4th ed.). Thousand Oaks, CA: Sage.
- Kulick, G., & Wright, R. (2008). The Impact of Grading on the Curve: A Simulation Analysis. *International Journal for the Scholarship of Teaching and Learning*, 2(2), n2.
- Kushman, J. W. (1997). Look who's talking now: Student views of learning in restructuring schools. Washington, DC: Office of Educational Research and Improvement, vol ED028257.
- Lambert, L. (2002). A framework for shared leadership. *Educational leadership*, 59(8), 37-40.
- Liu, X. (2008). *Measuring teachers' perceptions of grading practices: Does school level make a difference?* Paper presented at annual meeting of the Northeastern Educational Research Association, Rocky Hill, CT.

- Malouff, J. (2008). Grading bias. *College Teaching*, 56(3), 191-192.
- Maltese, A.V., Tai, R. H., & Fan, X. (2012). When is homework worth the time? Evaluating the association between homework and achievement in high school science and math. *High School Journal*, 96(1), 52-72.
- Marshall, C. and Rossman, G. (2016) *Designing Qualitative Research*. 6th Edition, SAGE, Thousand Oaks.
- Marzano, R. J. (2000). *Transforming Classroom Grading*. Alexandria, VA: ASCD.
- Marzano, R. J., Picerking, D. J., Pollock, J. (2001). *Classroom instruction that works: Research-based strategies for increasing student achievement*. Alexandria, VA: ASCD.
- Marzano, R. J. (2007). *The art and science of teaching: A comprehensive framework for effective instruction*. Alexandria, VA: ASCD.
- Marzano, R. J. (2010). *Formative assessment and standards-based grading*. Bloomington, IN: Marzano Research.
- McMillan, J. H. (2001). Secondary teachers' classroom assessment and grading practices. *Educational Measurement: Issues and Practice*, 20(1), 20-32.
- McMillan, J. (2003). Understanding and improving teachers' classroom assessment decision making: implications for theory and practice. *Educational Measurement: Issues and Practice* 34-43.
- McMillan, J. & Nash, S. (2000). *Teacher classroom assessment and grading practices*. Paper presented at the NCME annual meeting, New Orleans, LA.

- Merriam, S. B., & Tisdell, E. J. (2016). *Qualitative Research: A Guide to Design and Implementation (4th ed.)*. San Francisco, CA: Jossey Bass.
- Million, J. (1999). Restaurants, report cards, and reality. *NAESP Communicator*, 22(8), 5, 7.
- Mitra, D. L. (2018). Student voice in secondary schools: The possibility of deeper change. *Journal of Educational Administration*, 56(5), 473-487.
- Moss, C. M., & Brookhart, S. M. (2012). *Learning Targets: Helping Students Aim for Understanding in Today's Lesson*. Alexandria, VA: ASCD.
- Munoz, M., & Guskey, T. (2015). Standards-based grading and reporting will improve education. *The Phi Delta Kappan*, 96(7), 64-68.
- O'Connor, K. (2002). *How to grade for learning*. Thousand Oaks, CA: Corwin Press.
- O'Connor, K. (2009). *How to grade for learning, K-12 (3rd ed.)*. Thousand Oaks, CA: Corwin Press.
- O'Connor, K. (2011). *A repair kit for grading: Fifteen fixes for broken grades (2nd ed.)*. Boston: Pearson.
- O'Connor, K. (2017). A case for standards-based grading and reporting. *School Administrator*, 74(1), 24-28.
- O'Connor, K. (2018). *How to Grade for Learning (4th edition)*. Thousand Oaks, CA: Corwin Press.

- Olsen, B. & Buchanan, R. (2019). An investigation of teachers encouraged to reform grading practices in secondary schools. *American Educational Research Journal*, 56(5).
- Pattison, E., Grodsky, E., & Muller, C. (2013). Is the sky falling? Grade inflation and the signaling power of grades. *Educational Researcher*, 42(5), 259-265.
- Randall, J., & Engelhard, G. (2009). Examining Teacher Grades Using Rasch Measurement Theory. *Journal of Educational Measurement*, 46(1), 1-18.
- Randall, J. & Engelhard, G. (2009). Examining teacher grades using Rasch measurement theory. *Journal of Educational Measurement*, 46, 1-18.
- Randall, J. & Engelhard, G. (2010) Examining the grading practices of teachers. *Teaching and Teacher Education*, 26, 1372-1380.
- Reeves, D. B. (2004). The case against the zero. *Phi Delta Kappan*, 86(4), 325-225.
- Reeves, D. (2011). Elements of grading: A guide to effective practice. Bloomington, IN: Solution Tree Press.
- Reeves, D. B. (2011). Taking the grading conversation public. *Educational Leadership*, 69(3), 76-79.
- Reeves, D. B. (2008). Leading to change: Effective grading practices. *Educational Leadership*, 65(5), 85-87.
- Reeves, D., Jung, L. A., & O'Connor, K. (2017). What's worth fighting against in grading. *Educational Leadership*, 74(8), 42-45.

- Ross, J. A. & Kostuch, L. (2011). Consistency of report card grades and external assessments in a Canadian province, *Education Assessment, Evaluation, and Accountability*, 23, 159-180.
- Saldana, J. (2014). *Coding and analysis strategies*. In P. Leavy (Ed.), *The Oxford Handbook of Qualitative Research* (pp. 581-605). Oxford University Press.
- Saldana, J. (2016). *The coding manual for qualitative researchers (3rd ed.)*. London, UK: Sage.
- Schimmer, T. (2016). *Grading from the inside out: Bringing accuracy to student assessment through a standards-based mindset*. Bloomington, IN: Solution Tree Press.
- Scriffiny, P. (2008). Seven reasons for standards-based grading. *Association for Supervision & Curriculum Development*, 70-74.
- Selby, D., & Murphy, S. (1992). Graded or degraded: Perceptions of letter-grading for mainstreamed learning-disabled students. *British Columbia Journal of Special Education*, 16(1), 92-104.
- Stiggins, R. L. (2004). New assessment beliefs for a new school mission. *Phi Delta Kappan*, 9, 22-27.
- Stiggins, R., Frisbie, D., & Griswold, P. (1989). Inside high school grading practices: building a research agenda. *Educational Measurement: Issues and Practices*, 8(2), 5-11.
- Swan, G. M., Guskey, T. R., & Jung, L. A. (2014). Parents' and teachers' perceptions of standards-based and traditional report cards. *Educational Assessment, Evaluation and Accountability*, 26, 289-299.

- Thiele, Julie (2018) "Students' Perceptions of Grading Practices and Impacts on Mindset Development," *The Advocate*: Vol. 23: No. 5.
- Tomlinson, C. A. (2005). Grading and differentiation: Paradox or good practice. *Theory Into Practice*, 44(3), 262-269.
- Townsley, M., & Buckmiller, T. (2020). Losing As and Fs: What works for schools implementing standards-based grading? *Educational Considerations*, 46(1), 1-10.
- Vatterott, C. (2015). *Rethinking Grading*. Alexandria, VA: ASCD.
- Welsh, M. E., D'Agostino, J. V., & Kaniskan, R. (2013). Grading as a reform effort: Do standards-based grades converge with test scores? *Educational Measurement: Issues and Practice*, 32(2), 26-36.
- Westerberg, T. R. (2016). *Charting a Course to Standards-Based Grading*. Alexandria, VA: ASCD.
- Wiggins, G. (1996). *Honesty and fairness: Toward better grading and reporting*. In T. R. Guskey (Ed.), *Communicating student learning: 1996 Yearbook of the association for supervision and curriculum development* (pp. 141-176). Alexandria, VA: Association for Supervision and Curriculum Development.
- Wormeli, R. (2006). *Fair isn't always equal: Assessing and grading in the differentiated classroom*. Portland, ME: Stenhouse.
- Wormeli, R. (2018). *Fair isn't always equal (2nd edition): Assessment and grading in the differentiated classroom*. Portsmouth, NH: Stenhouse.

Yin, R. K. (2018). *Case study research: Design and method (6th Ed.)*. Thousand Oaks, CA: Sage.

Zoeckler, L. (2007). Moral Aspects of Grading: A Study of High School English Teachers' Perceptions. *American Secondary Education*, 35(2), 83-102.

Appendix A

Individual Student Interview Questions

The one-on-one individual student interview process will start with a welcoming introduction, an offer of snacks and beverages, and reiteration of the purpose of the study and ensure an overall understanding that this is an opportunity for the student to confidentially provide their opinions and for their perspectives to be heard.

1. (Introduction) Please introduce yourself and tell me a little bit about yourself.
 - a. How long have you been attending school here?
 - b. What have you enjoyed most about your high school career?
 - c. What activities and co-curriculars have been active in?
 - d. Tell me about your school. How might you describe it to a friend or someone who is considering attending school here?
2. Please tell me about your new grading system.
 - a. What can you share about grading practices in your school? How does it work?
 - b. How has what you described different from the previous grading practices/system you may have experienced?
 - c. Are there any similarities between the past and current practices based on your experiences?
 - d. Are there any specific grading practices that have carried over from the previous system?
 - e. Based on your experience, what are the benefits of the newer grading practices? What are the challenges?
3. Next, let's discuss how you perceive grading worked and felt in the previous system versus more recently in the newer grading system.
 - a. Does one system feel easier to understand and/or navigate when it comes to overall work than the other?
 - b. Does one system allow for you to change or improve your grades, including final grades, more than the other?

- c. Does one system seem to have better communication from teachers than the other?
 - d. What overall thoughts about the difference in systems do you have?
 - e. Which system do you prefer? Why?
4. Let's dialogue regarding how the grading systems relate to learning.
 - a. Which grading system do you think reflects your learning?
 - b. Why do you feel that way?
5. If you could be in charge of creating or changing a grading system in your school for current and future students, what might that look and feel like?
 - a. How might it be better than your current or previous grading system?
 - b. How might it look and feel different from what you have experienced throughout your high school career?
6. (Ending) Is there anything that has come to mind during our discussion today that you haven't had a chance to add or share?

Appendix B

Semi-structured Interview Questions for School Administrators

1. Please tell me a little bit about yourself.
 - a. What have you enjoyed most about your current role in your school district?
 - b. What are the most rewarding aspects of being a school administrator?
2. Please tell me about the grading experiences you have had in public education throughout your life.
 - a. What was memorable to you when you were a student?
 - b. What stands out to you as memorable from your administrator role?
3. Tell me about your journey moving from a traditional grading system to that of your current reformed grading system.
 - a. Did students play any role in this process?
 - b. Why or why not?
 - c. If students did play a role, to what degree does your district shift/modify the system based on student voice or feedback?
 - d. Will you consider using student input/feedback with your continued and future implementation?
4. Next let's talk about the transition between both systems.
 - a. What do you believe the transition process itself was like for students who were the first way of learners to experience both the former and the new reformed grading process?
 - b. How do you believe this will be different for future students?
 - c. How do you believe students initially perceived the shift in grading practices? Has that perception changed at all up to today?
5. Let's talk about student experiences with the grading system.
 - a. What means do you consider when trying to gauge student perceptions about their experiences in school?
 - b. What means do you consider when trying to gauge student perceptions about their experiences in school, specifically this shift in grading practices?

6. If you could go back in time and change one or two things about your implementation of and/or communication of the new grading system and practices, what might that be?
7. Is there anything else you would like to share regarding the evolution of grading practices in your building?

Appendix C

Semi-structured Interview Questions for Teachers

1. Please tell me a little bit about yourself.
 - a. What have you enjoyed most about your current role in your school district?
 - b. What are the most rewarding aspects of being a teacher?
2. Please tell me about the grading experiences you have had in public education throughout your life.
 - a. What was memorable to you when you were a student?
 - b. What stands out to you as memorable from your role as a teacher?
 - c. Do you believe any of your personal experiences as a student informed your practices/beliefs as an educator?
3. Let's talk about what other factors have informed your grading approach as an educator.
 - a. For example, were there any specific professional development or professional learning experiences about grading that stand out as memorable to you?
 - b. Any specific instruction or training when you were in your undergraduate training program?
 - c. Any district, building, or team grading policy/procedures that were a factor?
4. Next, let's talk about the fairly recent district shift in grading practices.
 - a. Please tell me about your experiences moving from a traditional grading system to that of your current reformed grading system in your school building?
 - b. Were you initially on board with the changes?
 - c. Do you believe in these changes?
 - d. How do you perceive students have been impacted by these changes?
 - e. Do you believe students
5. Let's talk about your current grading practices. Do the grading practices in your classroom include any of the following student factors, such as:
 - a. Responsibility;
 - b. Respect;
 - c. Timeliness

- d. Adherence/compliance to behavioral expectations?
6. Let's talk about the grading transition process your district recently went through.
 - a. What do you believe the transition process itself was like for students who were the first wave of learners to experience both the former and the new reformed grading process?
 - b. How do you believe this will be different for future students?
 - c. How do you believe students initially perceived the shift in grading practices?
 - d. Has that perception changed at all up to today?
7. Now let's talk about the grading practices themselves.
 - a. Do you believe your grading practices are an accurate means of reporting what students know and/or are able to do as it relates to the intended curriculum you deliver?
 - b. Why or why not?
8. If you could go back in time and change one or two things about your implementation of and/or communication of the new grading system and practices, what might that be?
9. Is there anything else you would like to share regarding the evolution of grading practices in your building?

Appendix E

Observational Protocol of School Building/Team Meetings

Date:

Time:

Location:

Conversations/discussions specific to grading components, strategies, implementation and outcomes	Occurred During Observation	Notes
Aspects of traditional grading		
Formative Assessments		
Homework		
Summative assessments		
Percentages and percentiles		
Letter grade		
Teacher perceptions and/or beliefs		
Aspects of teacher autonomy		
References to other counter-indicative practices (e.g., punishment; curve; inflation, etc.)		
Aspects of grading reform Components		

Determined standards		
Learning targets and evidence of learning		
Practice/mastery work without penalty		
Teacher perceptions and/or beliefs		
References to grading practices aligned to progress, product, and/or process		
Aspects of student voice		
Informal feedback from students		
Formal feedback from students (e.g., surveys; exit slips)		
Evidence of student perceptions and beliefs		