

Gardner-Webb University

Digital Commons @ Gardner-Webb University

Education Dissertations and Projects

School of Education

2019

A Grounded Theory Exploration of Teacher Grading Practices

Ingrid Rockhead

Follow this and additional works at: https://digitalcommons.gardner-webb.edu/education_etd



Part of the [Educational Assessment, Evaluation, and Research Commons](#), [Educational Leadership Commons](#), and the [Educational Methods Commons](#)

A GROUNDED THEORY EXPLORATION OF TEACHER GRADING PRACTICES

By
Ingrid Rockhead

A Dissertation Submitted to the
Gardner-Webb University School of Education
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Education

Gardner-Webb University
2019

Approval Page

This dissertation was submitted by Ingrid Rockhead under the direction of the persons listed below. It was submitted to the Gardner-Webb University School of Education and approved in partial fulfillment of the requirements for the degree of Doctor of Education at Gardner-Webb University.

Phillip Rapp, Ed.D.
Committee Chair

Date

Steve Laws, Ed.D.
Committee Member

Date

David Shellman, Ed.D.
Committee Member

Date

Prince Bull, Ph.D.
Dean of the School of Education

Date

Dedication

To my family and all those who stood behind me on this long adventure. Thank you for all your support. I could not have made it without you!

Acknowledgements

This work would not have been possible without the support and generosity of many people who contributed to the work presented here.

I am especially indebted to all the members of my Dissertation Committee, Dr. Phillip Rapp, Dr. Shellman, and Dr. Laws. Each member has provided me with extensive professional guidance and has taught me much about qualitative studies and research in general. Special mention goes to my dissertation chair, Dr. Phillip Rapp, for his expertise, helpful comments, and overall confidence in my ability to complete this project.

I would also like to thank Dr. Thomas Guskey and Ken O'Connor, two authors and experts on the topic of grading, whose work and words both directly and indirectly helped to shape this project. During an email exchange with Dr. Guskey, he expressed hope that his book, *On Your Mark*, would help with the planning of my project, and then stated, "But it must be yours!" It was those final five words that stuck with me the most and which helped to shape what would happen next. Those words prompted me to think differently and gave me new confidence to design the study through a different pair of lenses, my own. Those words stirred me into action and led to me contacting well-known author and grading expert Mr. Ken O'Connor. After numerous email exchanges and a long telephone conversation, Mr. O'Connor became one of the main contributors to the study presented here. Thank you!

I would like to thank members of my family, the most important people in my world and to whom this work is dedicated. Most importantly, I would like to thank my parents, whose love and influence remain a constant presence in all my pursuits. Thank

you for supporting me in the pursuit of this project. Special thanks to my brother John for his unwavering support and confidence in my ability to get the job done. Finally, but not least of all, I would like to thank my loving and supportive husband, Garey, for his never-ending inspiration!

Abstract

A GROUNDED THEORY EXPLORATION OF TEACHER GRADING PRACTICES.

Rockhead, Ingrid, 2019: Dissertation, Gardner-Webb University.

An examination of the history of grading reveals that many of the problems associated with current traditional grading practices are not new. Despite impassioned pleas from researchers and grading experts, practices such as assigning zeros on a percentage grading scale, assigning grades for practice such as homework and classwork while students are still learning, using the mean to report on students' overall grades, and not allowing reassessments continue to occur. The purpose of this qualitative study was to describe and report on Title I teachers' grading practices and explain using grounded theory methods, the influences on those practices. This study sought to explore the various influences on teacher grading practices and understand the causes for teachers' reluctance to modernize their grading practices. The study found that traditional grading practices are still commonplace in classrooms and that these practices are heavily influenced by the existing district and school policies that are in place.

Keywords: teacher grading practices, school/district grading policies/guidelines, grounded theory study, qualitative study

Table of Contents

	Page
Chapter 1: Introduction	1
Statement of the Problem	3
Theoretical Framework	4
Standards-Based Grading	5
Purpose of the Study	9
Significance of the Study	11
Deficiency	12
Definition of Terms	12
Research Questions	14
Chapter 2: Review of Literature	15
Home School Communication	16
History of Grading in the United States	18
Reliability of Grading	25
The Case Against Percentage Grading Scales	29
Purpose of Grading	35
Feedback and Standards-Based Grading	37
How Effective is Feedback?	41
Standards-Based Grading	43
Learning in the Standards-Based Classroom	43
Learning Targets and Feedback	44
Fixes for Broken Grades	49
Why Grading Reform Frequently Fails (Lessons Learned)	51
Summary	53
Chapter 3: Methodology	55
Introduction	55
Research Design	57
Participants	59
Rationale for Target Population	59
Purposeful Sampling	60
Theoretical Sampling	62
Instrumentation	63
Procedures	66
Delimitations and Limitations	73
Summary	74
Chapter 4: Results	76
Introduction	76
Participants	78
Categories	80
Research Questions	80
Research Question 1	81
The “E” Grade	82
The Zero Grade	88
Homework	90
Reassessment/Retake	96

Research Question 2	99
Policies (The Policies and Practices that Bind)	100
Expectations and Intent	101
Grade Composition	105
Documentation/Communication	107
School Policies.....	108
Research Question 3	110
A Long Talk with Grading Expert and Author Ken O'Connor	110
An Evaluation of One District's Grading Policy (The Policy, Problems, and Opportunities for Change).....	111
Opportunities for Change.....	116
Reflections from Participants.....	121
Axial Coding.....	122
Propositions.....	124
Selective Coding/Visual Model	126
Summary of Findings.....	127
Chapter 5: Conclusions	128
Introduction.....	128
Summary of Findings.....	129
Implications.....	142
Recommendations	143
Recommendations for Future Studies	144
References	150
Appendices	
A Consent to Participate	159
B Teacher Interview Protocol	162
C Principal Interview/Questionnaire	166
D District Leader Interview Protocol.....	170
E Permission to Conduct Study.....	174
F Participant Profile	176
G Author/Grading Expert Interview	180
H District Grading Policy	183
Tables	
1 Influences on Learning – Hattie.....	40
2 Summary of Effect Sizes Relating to Types of Feedback	42
3 A Comparison of Traditional and Standards-Based Grade Books	48
4 Open Coding Analysis of Grading Practices	82
5 Properties and Dimensions of Grading Practices.....	99
6 Open Coding Analysis of Policies	101
7 Required Grading Documentation	108
8 Evaluation of the Grading Documentation Expectations by Ken O'Connor.....	112
9 Ken O'Connor's Summary Evaluation of the District's Homework Policy.....	117
10 Ken O'Connor's Summary Evaluation of the District's Reteach/Retest Policy..	120
Figures	
1 Data Collection Activities.....	66
2 Axial Coding Paradigm.....	123

3	Visual Model Showing the Influences on Teacher Grading Practices	126
4	Distortion of Achievement Grade	135
5	The Overall Grade – A Confluence of Factors	137

Chapter 1: Introduction

In an age when educators at the state and local level are building accountability systems based on content and performance standards, grading practices have remained largely untouched. Practices that can be traced back to the late 1800s are still the most common in our schools. (Trumbull & Farr, 2000, p. xix)

The system of grades currently being used in most schools today is largely unchanged since its inception in the United States more than a century ago (Marzano, 2000). Guskey (2015) reported that “many aspects of grading and reporting reflect traditions that have been a part of our educational system since the time our great-grandparents were in school” (p. 2). Furthermore, this current and traditional grading system, despite being in use for so long, has drawn heavy criticism by many experts in the measurement community and is regarded as being unreliable, inaccurate, and damaging to students (Guskey, 2011; Vatterott, 2015).

At present, many of the questions and criticisms about grading practices are not new and have all been forcefully presented years ago by several researchers and grading experts (Finkelstein, 1913; Kirschenbaum, Simon, & Napier, 1971; Kohn, 2011). To illustrate, Finkelstein (1913) posed the following questions years ago:

It is within the last decade that serious attention has been paid to such queries as: What should the mark really represent? Should the mark be based on ability or performance, or even upon zeal or enthusiasm? ... Is it possible, ... to increase the fairness and reliability of marks? Do students tend to secure the same standing under different teachers in the same school? (p. 5)

In addition, nearly 100 years later, Stiggins (2000) posed several uncomfortable

questions which bare great similarity to the ones raised by Finkelstein (1913). They also serve to highlight the problems with traditional grading practices that have been known for some time.

When we confront the challenge of implementing sound grading practices, for example, we must make many crucial decisions. First, we must decide the purpose of those grades. Will they serve to motivate, or to communicate? ... Or what if honest communication about low achievement is demoralizing for the unsuccessful student, causing that student to give up in hopelessness? ... Do we wish to share information about achievement, intelligence, effort, attitude, compliance with rules...? What happens when we pack all of these together in one grade and the message receiver is unable to sort out the contribution of our various ingredients and thus is unable to discern what the resulting grade really means? Can effective communication result? (Stiggins, 2000, p. x)

It is evident that the problems with grading and reporting have been known for decades (Guskey, 2015); and today, several prominent individuals in the field of education and research continue to present compelling arguments in support of grading reform (Brookhart, 2015; Guskey, 2011; Kohn, 2011; Marzano, 2000; O'Connor, 2011; Schimmer, 2016; Vatterott, 2015). One of the earlier and more outstanding examples is the study by Starch and Elliot (1912) which called into question the reliability of grades. Furthermore, nearly 100 years after Starch and Elliot's (1912, 1913) groundbreaking research, Brimi (2011) replicated this study to further explore the reliability of grading and how grading practices have evolved over time. In this more recent study, Brimi concluded that very little has changed and that grading today continues to be subjective

and unreliable. These studies continue to reveal that many of the problems associated with grading not only existed years ago but continue to be perpetuated today.

Statement of the Problem

There is growing evidence today that reveals that a discrepancy exists between the actual grading practices currently being employed by educators and those practices recommended by grading experts. To illustrate, many teachers still act on the mistaken belief that the use of zeroes on a 100-point scale is an appropriate consequence for students who fail to turn in work (Frisbie & Waltman, 1992; Reeves, 2004, 2008). Several teachers today still disallow retakes on assignments that are poorly done (Wormeli, 2011). Rather than offering feedback, many teachers continue to penalize students by grading their work while the learning is still underway (Vatterott, 2015). The practice of averaging various sources of evidence such as homework, quizzes, tests, projects, classwork, work habits, and effort and subsequently reporting this as one overall grade per subject at the end of the grading period is still prevalent in many of our schools today (Guskey, 2011, 2015; Schimmer, 2016). Other researchers who have weighed in on the subject of the single grade purport that in order for a grading system to be more effective, the “overall” grade should be eliminated (Marzano & Heflebower, 2011), since “a single grade hides more than it reveals” (Wiggins, 1996, p. 145).

The shortcomings of this practice are also summed up by Schimmer (2016) who recently noted,

For decades, teachers have relied on the mean average to calculate grades, yet the mean rarely reflects what students know in the end, as it is vulnerable to extreme or atypical results. The mean can be *mean* by not giving students full credit for

where they are along their continuum, even when the calculation is spot on. (p. 61)

Despite the knowledge we have accumulated regarding what constitutes good grading habits, many damaging and outdated practices still exist.

This study sought to explore the various influences on teacher grading practices and understand the causes for teacher reluctance to modernize their grading practices.

Theoretical Framework

“Education improvement efforts over the past two decades have focused primarily on articulating standards for student learning” (Guskey, 2011, p. 17). This focus on standards has its roots in the controversial but highly influential 1983 report known as *A Nation at Risk* (National Commission on Excellence in Education, 1983). Education historian Diane Ravitch (2010) recounted that the report was an appeal for the public’s attention to the fundamental issues in education at the time:

It warned that the nation would be harmed economically and socially unless education was dramatically improved for *all* children.... To that end, the report recommended stronger high school graduation requirements; higher standards for academic performance and student conduct; more time devoted to instruction and homework; and higher standards for entry into the teaching profession and better salaries for teachers. (p. 25)

In response to some of the concerns raised in *A Nation at Risk* regarding the quality of education, most states worked assiduously to develop and define standards for most of the key content areas taught in the U.S. in the years following its publication. This was a welcomed move by most educators. According to Guskey (2009),

The release of the first set of standards by the National Council of Teachers of Mathematics in 1989 was greeted with unprecedented optimism.... Soon thereafter, other professional organizations followed suit. The National Council for the Social Studies (1994), National Academy of Science (1996), National Council of Teachers of English (1996), and the American Council on Teaching Foreign Languages (1996) all developed standards in their respective disciplines. States and provinces also took up the task and, today, nearly all have identified standards for student learning. (p. 1)

Consequently, in the years following the publication of *A Nation at Risk*, one of the most dominant issues in education has been related to establishing standards. The implementation of standards-based reforms not only provided much needed clarity on what students should learn but also presented a major challenge, that of grading and reporting on those standards (Guskey, 2009). Today, standards are now well established, and despite this, the problem with which educators are now faced is “the daunting task of how best to grade and report student learning in terms of those standards” (Guskey, 2001, p. 20). According to Schimmer (2016), it only stands to reason that teaching to standards should also imply grading and reporting on those very standards.

Standards-Based Grading

Several prominent individuals in education have proposed standards-based grading as the better alternative to the traditional grading system still commonly practiced today (Guskey, 2015; O’Connor, 2011; Vatterott, 2015). Just as the standards movement resulted in curriculum that was better aligned, McMillan (2009) explained that standards-based grading will result in greater consistency with our grading practices. Researchers

argue that standards-based grading provides a more accurate picture of student learning since it is more focused on student mastery of content material (Guskey, 2001; Schimmer, 2016; Vatterott, 2015).

Vatterott (2015) regarded standards-based grading as a change that moves beyond grading itself into the process of achieving a “complete overhaul of the teaching learning process” (p. 26).

Researcher, professor, and author Cathy Vatterott (2015), who is a strong supporter of standards-based grading, raised concerns which bare noticeable similarity to those expressed in *A Nation at Risk*. One of the indicators of the risk outlined in the report was that a large majority of 17-year-old students lacked the necessary intellectual skills that would enable them to draw inferences, write persuasive essays, and solve multistep math problems. Similarly, Vatterott also wrote about her college freshmen and sophomores:

They often don't fully comprehend how to analyze and synthesize. They seem to be stuck in the mode of “just tell me what you want.” Many of them are terrible writers, unable to express their thoughts clearly and intelligently. I see firsthand the damage we have done, and how we have handicapped them for college by giving grades that don't reflect learning. I am not surprised by how many drop out. We have the opportunity to change that.... Standards-based Grading has the potential to restore integrity to the grading process. It can and will change our students' futures. (pp. 1-2)

In addition, Vatterott explained,

We know that something is wrong with grades. Every day we see the mismatch –

on one hand, the stellar performance on standardized tests from B and C students (thus labeled “underachievers”), and on the other, poor performance on standardized tests from straight A students. We know that many students leave high school with high grade point averages yet struggle academically in college. (p. 5)

Vatterott (2015) argued that such discrepancies in grade point averages and student performance on standardized tests only perpetuate the current high school and college dropout rate. Vatterott also cited other sources of evidence that not only point to the need for educational reform but also support her argument for grading reform. To illustrate, she cites U.S. student performance on international assessments like the Programme for International Student Assessment (PISA). This international assessment is one of the most challenging and commonly used international tests which is used to measure and compare the performance for 15-year-old students in over 65 countries around the world. PISA, which is coordinated by the Organization for Economic Cooperation and Development (OECD), has been administered every 3 years since 2000. Results from PISA in 2012 revealed that students in the U.S. scored much lower than other developed countries and were ranked 20th in reading, 23rd in science, and 30th in math among the participating countries. (U.S. Department of Education, 2015, as cited in Vatterott, 2015).

Vatterott (2015) avouched that the reform of this one educational practice, grading, could potentially influence major changes and reform in K-12 education. In addition, she purported that many of the educational ills observed today can be alleviated through grading reform.

According to Muñoz and Guskey (2015), the journey to accomplishing sound grading practices begins with taking steps to ensure that grades are meaningful. This can only be realized if educators and other stakeholders first examine and acknowledge the many ways in which their beliefs can and have influenced their practice (Brookhart, 2011; Vatterott, 2015). When exams, quizzes, projects, reports, homework, punctuality in turning in assignments, work habits, and effort are merged together and reported as a single grade, it is difficult to interpret such grades correctly and meaningfully. This is the case with traditional grading systems. In order to address this problem, many researchers have proposed standards-based grading as a possible solution to reforming such damaging grading practices (Guskey, 2011; Marzano & Heflebower, 2011; O'Connor, 2011; Vatterott, 2015).

Despite the strong focus in recent years on standards and assessment of student proficiency on those standards, many agree that “there remains one arena where few educational leaders and reformers have ventured: classroom grading” (Cox 2011, p. 67). Schimmer (2016) explained that “realigning grading practices to the now well-established standards-based instructional paradigm represents the final step of the assessment revolution” (p. 15).

According to Reeves (2016), “Changes in grading systems are often the last things on the agenda for school reform” (p. 19). Guskey (2011, 2015) similarly contended that grading and reporting on student achievement is the one element still missing from major education reform efforts.

Notwithstanding, several leading researchers continue to forcefully argue for grading reform (Cizek, 1996; Clymer & Wiliam, 2007; Guskey, 2011, 2015; Marzano &

Heflebower, 2011; Muñoz & Guskey, 2015; Reeves, 2011; Vatterott, 2015; Winger, 2005). They have presented impassioned arguments against the use of zeroes; norm-based grading; grades to motivate students; grades as a means of differentiating students; and the use of one grade, encompassing nonacademic factors like effort and behavior, to sum up a student's achievement in a subject (Brookhart, 2011; Guskey & Bailey, 2001; Marzano & Heflebower, 2011; Schimmer, 2016; Vatterott, 2015). Despite the strong research base which currently exists on recommended grading practices, Guskey (2011) noted, "everyday teachers continue to combine aspects of students' achievement, attitude, responsibility, effort, and behavior into a single grade that's recorded on a report card – and no one questions it" (p. 19).

Schimmer (2016) contended that it is now time to break away from the tradition of grading and embrace the new grading paradigm:

It's time we pushed aside our old ways of thinking and took a fresh look at how we report student achievement. Developing a new grading paradigm is the necessary first step toward significant grading reform, but we can't make this breakthrough unless we are prepared to renounce our traditional approach to grading. This new paradigm is the standards-based mindset. (p. 3)

Purpose of the Study

Why has a system that is so widely criticized endured for such a long time? Guskey (2015) posed other salient questions which, to some degree, remain unanswered and which this study therefore seeks to explore more closely: "Why, if we've known about these problems for so long, have we not found a solution? With all that we have learned about education over the past hundred years, why have grading and reporting

continued essentially unchanged?” (p. 3). In response to these questions, Guskey (2015), posited that the reason our grading practices have remained unchanged is due to the grave consequences associated with grades. To illustrate, Guskey (2015) explained,

Grades largely determine whether or not students get promoted from one grade level to the next. They are used for determining honor roll status, membership in honor societies, and enrollment in advanced classes. High grades are required for admission to selective colleges and universities, and low grades typically are the first indicator of potential learning problems. (p. 3)

Additionally, Guskey (2015) theorized that educational leaders fear that since “the relationship between grades and these consequences is so powerful ... changing the way we grade could disrupt that relationship and confound crucial decision-making procedures” (p. 3).

The purpose of this qualitative study was to describe teacher current classroom grading practices and to explain the process by which teacher grading practices develop and evolve throughout their teaching careers. Special attention was given to understanding why teachers continue to embrace traditional grading practices. The study also sought to explore the various ways in which teachers communicate about student learning on the standards taught and offer feedback as students work toward mastery of those standards. Additionally, the study describes the school and district grading policies in order to determine the extent to which these policies influence teacher grading practices.

The data generated from this study could have practical application for the school district by informing district leaders and other stakeholders about possible professional

development needs of teachers and administrators relating to standards-based grading. The study could also help to reveal deficiencies in district grading policies that might be worthy of attention; therefore, a fundamental reason for this study was to inform decision makers about the current grading climate of schools in the district so they can identify needs and develop action plans to address those needs.

Guskey (2011) maintained that the task of challenging the ineffective traditional grading practices is not an easy one since they have been a part of most people's educational experiences for so long. In order for educational leaders to be more successful with tackling this problem, Guskey (2011) stated that leaders must be "familiar with the research on grading and what works best for students, so they can propose more meaningful policies and practices that support learning and enhance students' perceptions of themselves as learners" (p. 21). Guskey (2011) further contended that "leaders who have the courage to challenge the traditional approach and the conviction to press for thoughtful, positive reforms are likely to see remarkable results" (p. 21).

Significance of the Study

Grades continue to be one of the most common means of communication and measurement of student learning in schools (Cizek, 1996; Cox, 2011; Guskey, 2011). Grades also serve as the "primary indicator of how well students perform in school and is the basis for making important decisions about students" (Guskey, 2015, p. 3). To a great extent, grades "certify competence, permit graduation, advance students to the next grade, and predict future achievement" (Summers, 2009, p. 5). Grades are usually the first indicator of possible disabilities. Grades help determine honor roll status and also

admission to certain universities or colleges (Guskey, 2015). Notwithstanding, the common theme in the literature today and in previous years is the discrepancy between actual grading practices of teachers and those best practices which have been recommended by grading experts (Brookhart, 1993; Guskey, 2000; Stiggins, 2000; Vatterott, 2015).

According to Reeves (2016), if he could offer one suggestion that would “in less than a single year, reduce failure rates, improve discipline, and increase faculty morale” (p. 2), that suggestion, which is supported by research, would be for schools to improve grading practices. Some of the experts who echo similar sentiments are Schimmer (2016), Guskey (2015), Reeves (2012), O’Connor (2011), Marzano (2006), and Guskey and Bailey (2001). Reeves (2016) further attested that when schools take steps to “embrace effective grading practices and terminate toxic grading policies, student performance, motivation, and discipline improve” (p. 2).

Deficiency

There is a deficiency in specific legislation to guide teachers on grading practices. In many districts, grading practices tend to be predominantly guided by district policies. Consequently, “thoughtful policies, clearly communicated and uniformly applied, are the key to legally defensible grading practices” (McElligott & Brookhart, 2009, p. 67).

Definition of Terms

In order to provide a common understanding of what is meant by key terminology, this section provides definitions of key concepts discussed in this study.

Assessment. Vehicles for gathering information about student achievement or behavior (Marzano, 2000). Planned or serendipitous activities that provide information

about student understanding and skills in a specific measurement topic (Marzano, 2006). Similarly, Hattie and Timperley (2007) regarded assessment as activities that provide teachers and/or students with feedback information relating to the discrepancy between current status and learning goals.

Evaluation. The process of making judgments about the level of student understanding or performance (Marzano, 2000).

Measurement. The assignment of marks based on an explicit set of rules (Marzano, 2000).

Score. The number or letter assigned via a process of measurement. The term mark and score are commonly used synonymously.

Grade. The number or letter reported at the end of a set period of time as a summary statement of student performance (Brookhart, 2015) or the marks on individual pieces of student work that make up that summary (Brookhart, 1994).

Marks. As defined by Starch and Elliot (1912), the marks or grades attached to a pupil's work are the tangible measures of the result of his attainments (p. 442).

Grading. Teacher evaluations, formative or summative, of student performance.

Grade inflation. An increase in grades without simultaneous increase in student achievement.

Grading practices. The methods teachers use to determine student grades, including the factors contributing to the formation of grades (Bailey, 2012). Grading practices also refer to all the methods teachers use to rate student work or provide feedback on their academic coursework (Dyb, 2011).

Standards/objectives. These are clear statements that define what students at

each grade level should learn and be able to do.

Standards-based grading practices. Guskey and Jung (2009) defined standards-based grading practices as using a report card that “includes grades or marks based on carefully articulated learning standards in each subject area” (p. 1).

Traditional grading practices. According to Hooper and Cowell (2014), traditional grading practices include assigning “points on various types of assignments and assessments throughout a grading period and a teacher averaging those points on a 100-point scale to determine a student’s overall grade” (p. 59).

Research Questions

The specific research questions for this study are

1. What are the grading practices currently used by teachers?
2. What factors influence the development and evolution of teacher grading practices?
3. What changes (if any) might be implemented in current grading practices which would provide a better measure of assessing student achievement?

Chapter 2: Review of Literature

This chapter provides a framework for this study by reviewing the literature relevant to the research topic. Sections of the literature review are structured around themes relating to assessments. In particular, the study focuses on how classroom assessments are formulated and reported and their relationship to student learning/achievement.

The first section introduces the topic of grades as a form of communication. Following this, the literature on grades and the history of grading is reviewed; subsequently, the discrepancies between actual and recommended practices of teachers are highlighted. In the later sections, the research on feedback is presented which culminates into the final segment of the chapter where the case for standards-based grading is presented.

Stiggins (2000) expressed the following:

If schools are to be effective – that is, if instruction is to result in maximum student learning – then educators must *communicate* effectively about student achievement. Important decision makers need access to dependable information about student achievement in order to do their jobs. Without it, how can they diagnose student needs, allocate resources, pace instruction, evaluate the effectiveness of instructional interventions, or provide feedback to students? (p. ix)

These statements by Stiggins appropriately capture the most important objective of schools, student learning. His statements also outlined an important action, frequently overlooked, which needs to take place in order to enhance this learning – effective

communication (Trumbull & Farr, 2000).

According to Guskey (1996), “The primary goal of grading and reporting is communication” (p. 3). Additionally, for over 100 years, teachers in the United States have been using grades as a means of communicating student achievement (Marzano, 2000); however, throughout much of this time, education researchers have also remained strongly critical of its use. Despite the prevalent use of grades, many experts argue that grades reflect one of the most unreliable means of measurement in education and an area where few educational leaders dare to venture (Cizek, 1996; Cox, 2011; Guskey, 2011). According to Cizek (1996), no matter the form in which grades come, they are simply “primitive tools for doing the job” (p. 104). As a means of communication, he compared the grading system as being as ineffective as two tin cans connected by a string. Similarly, in reference to grades, Kohn (2011) regarded them as being “a relic from a less enlightened age” (p. 28).

Home School Communication

Waltman and Frisbie (1994) carried out a study to determine if parents interpret their children’s report card grades in the same way as the teachers who assigned the grades. The participants in the study consisted of 16 teachers and 285 parents of their students. Questionnaires pertaining to the meaning of report card grades were completed by the participants. From this study, the authors concluded that the results overwhelmingly reflected that report cards failed to communicate the teacher’s intended meaning to parents.

More recently, Webber and Wilson (2012) carried out a related investigation to determine the nature of the communication parents desired. In this effort, they sought to

discover if parents were satisfied with report card grades or if they wanted more detailed descriptions of their child's progress and challenges. In order to make this determination, parents of students in one teacher's classroom were invited to participate in an hour-long interview about home school communications. Seven parents responded, and all were interviewed. Webber and Wilson made the following observations after the interviews: "Every parent insisted that access to student work, teacher comments, conversations with the teacher, and narrative descriptions of learning were most important" (p. 32); and "Parents want teachers who observe their children carefully, develop strong learning relationships and communicate meaningfully.... Parents are telling us what they want: fewer grades, more description, and more shared artifacts of teaching and learning" (p. 35).

Some of these observations made by Waltman and Frisbie (1994) and Webber and Wilson (2012) seem to echo statements by Marzano (2000) who, in the first chapter of his book *Transforming Classroom Grading*, explained that there is compelling evidence that indicates that grades are so imprecise that they no longer communicate valuable information about student learning. He further explained that the grading system today "is at least 100 years old and has little or no research to support its continuation" (Marzano, 2000, p. 13; Guskey, 2011). These statements are corroborated by Farr (2000) who also noted that traditional grading practices in the United States represent a part of our "educational history that has been almost impervious to change" (p. 3).

In order to fully appreciate the arguments made by the many individuals calling for grading reform, even to the extent of grade abolition, and in order to evaluate statements that the current grading system is largely unchanged and ineffective, it will be

helpful to first examine the history of grading in the United States, studies which have emerged over the years, and recommendations that have been proposed.

History of Grading in the United States

“The precise history of grading practices in American education is somewhat debatable” (Marzano, 2000, p. 11). To illustrate, Guskey (1996) reported that prior to 1850, grading and reporting were virtually unknown in U.S. schools (p. 14), whereas Marzano (2000) reported that much of the history of grading and reporting in the United States can be traced as far back as the 1700s. Despite such differences in historical accounts, “most historians agree on a number of significant events” (Marzano, 2000, p. 11), which are presented below.

In order to help place the reader in this time period and therefore paint a coherent picture of the education landscape during this time, a few historical milestones will be highlighted. These include the establishment of the first public school in the United States, Boston Latin School, founded in 1635, and the first institution of higher education, Harvard, established just over a year later in 1636. The college of William and Mary and Yale University were later founded in 1693 and 1701 respectively. In the beginning, according to Kirschenbaum et al. (1971), education took place mainly within the family (p. 47) and only became more widespread in the mid-19th century when mandatory school attendance laws began to emerge. Children of wealthy individuals received a higher quality of education for the purpose of helping them prepare for some of the earlier colleges which existed at that time like Harvard, William and Mary, and Yale (Kirschenbaum et al., 1971).

A review of most historical documents suggests that grading practices in the United States first developed in institutions of higher learning (Brookhart, 2004; Troxell, 2009). The earliest records of grading to differentiate students seemed to have been at Yale University in 1785 (Smallwood, 1935); however, education historian, Mark Durm (1993) noted that in all probability, some of the other colleges established before Yale undoubtedly used some method for evaluating students even if without a standard. Much of the history of grading in American colleges was documented by Smallwood (1935) in a book which later became her doctoral dissertation and is entitled *An historical study of examination and grading systems in early American universities*. One major source of information which contributed to Smallwood's work was the diary of Ezra Stiles, the then president of Yale in the late 18th century. Other sources consisted of original historical documents from Harvard, William and Mary, the University of Michigan, and Mount Holyoke (Durm, 1993; Troxell, 2009). Smallwood reported that "the first ... evidence of a real marking system is at Yale in 1785" (p. 42). This system, she explained, consisted mainly of descriptive adjectives (Smallwood, 1935). A footnote in Stiles's diary in 1785 made reference to an exam taken by 58 students, where there were "twenty *Optimi*, sixteen second *Optimi*, 12 *Inferiores*, (*Boni*), ten *Peiores*" (Smallwood, 1935, p. 42). It was customary at that time for students to remain in college for different lengths of time until they were deemed ready to pass a graduation exam, which was usually administered the day of graduation (Brookhart, 2004, p. 16; Smallwood, 1935). This record not only shows how students were evaluated after an examination but also appears to be "the very first collegiate 'grades' given in the United States" (Durm, 1993, p. 295). The record also seems to suggest that examinations during

that time were not the primary determinants of success or failure but instead may have been a formality since students had to have been previously deemed successful in order to be recommended for graduation. It also seems to suggest that some method of evaluation, other than the graduation exam, must have existed to allow them to be able to make the determination for which students would be ready for the graduation exam. Other than the footnotes found in Stiles's diary in 1785, there appears to have been a 28-year gap during which no other records of grades have been found. Records in 1813 revealed that Yale began giving feedback to students on a 4-point scale (Durm, 1993; Marzano, 2000). This appears to be the first record of a numerical scale being used (Brookhart, 2004). It was most likely Yale University that pioneered the 4.0 scale system which is still widely used today. Smallwood (1935) posited that this 4.0 scale could have been possibly linked to the four Latin terms *Optimi*, second *Optimi*, *Inferiores*, (*boni*), *Peiores* (p. 44), that were cited earlier in reference to exam grades. At this time, there was no record or evidence to suggest that letter grades were in use (Durm, 1993).

Following Yale's 4.0 system, other universities started to experiment with and employ various types of numerical scales and grading systems (Durm, 1993; Marzano, 2000). Durm (1993) and Smallwood (1935) offer accounts of what transpired in the years following 1813 as reflected below:

- In 1818, records reveal that William and Mary classified students using the categories No. 1, No. 2, No. 3, and No. 4. In this classification, No. 1 with the individual's name next to it represented the individual who was first in their class; No. 2 was the classification given to students who were orderly, correct, and attentive; No. 3 was a classification which indicated that the student made

very little improvement; and No. 4 meant the student learned little or nothing.

- In 1830, Harvard began to use its first numerical scale of 20, rather than 4; however, a few years later in 1837, some professors explored the use of the scale of 100. Records seem to indicate that this was short lived, as other faculty members quickly expressed a desire to revert to the 4.0 system that was originally being used.
- In 1850, William and Mary began to use its first numerical scale. In the years prior to 1850, the faculty had been using descriptive adjectives in reports that were sent home to parents.
- The University of Michigan explored several different systems including abolishing grades and replacing them with a pass-no pass system which it implemented in 1851 (Brookhart, 2004). In 1852, a plus symbol was used to represent a pass mark and a minus symbol was used to represent a fail mark. Later, in 1860, they added a conditional grade in addition to the plus sign; but shortly after this, they shifted to a 100-scale system. Seven years later in 1867, the university adjusted its scale yet again and began using a P to represent a passing grade, a C to represent conditional, and an A to represent absent (Durm, 1993, p. 296).

In the middle of the 19th century, while institutions of higher learning continued paving the way for other schools by experimenting with various marking systems and establishing and raising standards for awarding college degrees, government support for elementary schools had just started to gain popularity (Kirschenbaum et al., 1971). This was largely due to the passage of compulsory school attendance laws (Guskey, 1996), the

first of which was passed in Massachusetts in 1852.

It is important to recall that prior to 1852, few students had the privilege of attending school except those who were from wealthy families. Elementary schools during this period typically consisted of one-room schoolhouses with students of all ages grouped together with one teacher (Guskey, 1996; Kirschenbaum et al., 1971). Student progress was usually reported orally to parents through home visits (Guskey & Bailey, 2001, as cited in Troxell, 2009). In addition, many students rarely enrolled in schools past the elementary level, except for children of the wealthy, who typically had private tutors or attended special schools for the purpose of preparing them for acceptance to colleges like Harvard, William and Mary, and Yale. Also, during this time period, exams were given as a way of testing student knowledge, not for giving grades. Kirschenbaum et al. (1971) clarified this: “The purpose of this testing was two-fold: it demonstrated the student’s progress, and it gave the teacher a clearer indication of what subjects required additional instruction to enable the student to handle the work required in college” (p. 50).

As more states began passing compulsory school attendance laws, enrollment in elementary schools and high schools gradually increased (Guskey, 1996; Kirschenbaum et al., 1971). Consequently, these one-room schoolhouses were eventually transformed to schools which grouped students in grade levels based on their age. Reports from Kirschenbaum et al. (1971) indicated that between 1870 and 1910, the total number of students in elementary and high school in the U.S. increased from 6,871,000 to 17,813,000. As a result of these changes, curriculum and instruction gradually evolved and became more specific to subject and age level. This increase in the number of

students also led to the creation of new ideas in curriculum and instruction. Among these new ideas was the use of progress evaluations of student work. This would entail writing down and differentiating between the skills each student mastered and those for which more work was needed. This was done as a means of determining which student was ready to move to the next grade level (Guskey, 1996, 2013; Vatterott, 2015).

Furthermore, this increase in the number of students at the secondary level prompted educators to search for grading methods that were less burdensome than the traditional descriptive reports. As a consequence, several of these schools also began examining and adopting new grading techniques that were being used in colleges (Trumbull & Farr, 2000; Vatterott, 2015). In an effort to manage the increase in student population, teacher use of detailed narrative reports based on student work became less frequent; and instead, what became more commonplace was the use of “single numerical (or letter) grades based on some quantification of learning, most often a percentage” (Farr, 2000, p. 4). Farr (2000) noted that this “shift from the use of narratives to a more reductionist approach” (p. 4) is one of the more interesting aspects of the history of grading in the United States.

While compulsory school attendance laws sparked major changes in elementary and high schools, institutions of higher learning continued to experiment with various grading systems. Two such colleges which prompted shifts from the common narrative reports to a more quantitative approach were Harvard and Mount Holyoke. A summary of the way in which this shift unfolded is outlined below.

In 1877, Harvard transitioned from a 20-point grading system to a system where students were classified into six divisions (Marzano, 2000):

- Division 1: 90 or more on a scale of 100
- Division 2: 75-90
- Division 3: 60-74
- Division 4: 50-59
- Division 5: 40-49
- Division 6: below 40

Harvard's division system is believed to have been the precursor to the letter grades still in use today. This reasoning is supported by the records of 1897 which reveal that Mount Holyoke adopted the use of letter grades, combined with adjectives and percentages, a few years after Harvard's division system was implemented. This is also illustrated below:

- A: Excellent = equivalent to percents 95-100
- B: Good = equivalent to percents 85-94
- C: Fair = equivalent to percents 76-84
- D: Passed = barely equivalent to percent 75
- E: Failed = below 75

By the turn of the century, large increases in school populations sparked various changes mainly in high schools and colleges. Elementary schools continued to use written descriptions as grades, but high schools gradually employed the use of percentages or similar markings to document student achievement in various subject areas. Additionally, as the number of students graduating from high school increased, so did the number of college applicants. Over time, colleges needed a way to screen their applicants and high school percentage grades provided a convenient means to do so. This

new system of grading soon became the standard used by colleges to sort and rank its applicants. This percentage grading system adopted by Mount Holyoke in 1897 marked the beginning of the 100-year tradition of grading and reporting as we know it today (Kirschenbaum et al., 1971; Marzano, 2000, p. 11).

Reliability of Grading

It was not long after this percentage grading system gained popularity before educators started to question its reliability. Emerging from this debate was a powerful and influential study carried out by Starch and Elliot (1912) to examine the reliability of grading by English teachers in a school district. The purpose of the study was to determine the extent to which the personal values and expectations of teachers influence their grading (Guskey, 1996, p. 14).

In the original study by Starch and Elliot (1912), copies were made of two English exam papers written by two high school students at the end of their first year in high school. These copies were then distributed to over 200 schools with the intent of having them graded by, to the extent possible, teachers who had experience teaching first year high school English. For this study, 200 schools received copies of the work. Of this total, 142 schools returned the papers graded. The papers were graded on a scale of 0-100 and the passing score was 75%. The range of scores on each paper was alarming. For one of the two papers, the scores ranged from 64-98; and for the other paper, the scores ranged from 50-97. The researchers found that teachers were very subjective in their grading practices and that some focused on neatness, punctuation, elements of grammar, and style; while others valued the clarity of the writing (Guskey, 1996; Vatterott, 2015).

The study was criticized by many who claimed that “good writing is a highly subjective judgement” (Guskey, 1996, p. 15) and that English teachers are more likely to be subjective in their grading. In response to this criticism, Starch and Elliot repeated the study the following year in 1913, but this time they used geometry papers instead of English papers. The result of their study showed even greater variations in grades. Of the 200 papers that were distributed, 138 were returned graded with scores on one paper ranging from 28-95. The results showed that some teachers deducted points for incorrect answers; while others factored in neatness, form, and even spelling as part of the grade.

As a consequence of the findings from Starch and Elliot (1912), for a brief time period, there was a gradual shift away from the use of percentage grades to other grading scales with fewer categories. Today, the practice of percentage grading scales, which emerged in the 1880s from Harvard’s division system, is still in use but continues to be challenged by measurement experts. This groundbreaking study by Starch and Elliot (1912) demonstrated that percentage grades were unreliable and inaccurate. The study also revealed that there were wide variations in grading practices among teachers. As Goodwin (2011) reported in just over a decade following Harvard’s new approach to grading, even the professors began to note the shortcomings of the grading scale, complaining that As and Bs seemed to be issued too easily. Consequently, many educators gradually shifted from percentage grades to other scales with fewer categories. For a short time period following this shift to the use of fewer score categories, grades were more consistent; however, in the 1990s, as the use of electronic grading software programs increased in education, there was a corresponding resurgence of percentage grades (Guskey, 2013).

As percentage grading made a comeback, so did the criticisms from grading experts questioning its reliability. Nearly 100 years following the groundbreaking research by Starch and Elliott (1912), familiar concerns emerged regarding the reliability of percentage grades. As a result, the original research was later replicated by Brimi (2011) to further explore the reliability of grading and how grading practices evolved over time.

Some of the questions which influenced Brimi's (2011) work and the study he conducted were as follows: Do grading scales affect teachers grading/perception of the work? After 100 years following the work of Daniel Starch and Elliot, have the subjective elements of grading diminished? And would teacher grading of the exact English paper differ as greatly today as their 1912 predecessors?

One difference between the original study and that conducted by Brimi (2011) is that Brimi's study was focused on the reliability of grading by English teachers within one school district. Another difference in Brimi's study is that teachers were trained to use a grading system known as the NWREL's 6+1 traits of writing. This was unlike Starch and Elliot's (1912, 1913) study which used data from over 200 schools where teachers did not use any specific grading system.

In this single school district where the study was conducted, all teachers were trained on how to use the NWREL 6+1 model to assess writing. This training took place over a 2-day period in the summer of 2007. In the spring of 2008, the same teachers participated in a follow-up session at their schools. During this follow-up session, the researcher presented each teacher with an identical copy of a student paper to grade. The researcher had graded the paper too and had assigned it a grade of 83%, or a high C.

Participants in this study were asked to grade the paper on a 100-point scale using the 6+1 model to assess the work. They were also instructed to grade the work individually and without consulting other teachers. The graded work was then collected and coded. The results reflected that of the 73 papers that were scored, the range of scores was 43 points and the paper received 30 different scores (Brimi, 2011).

The study also showed that despite the training teachers received on a specific grading method, participants awarded grades on a student's work that was as varied as those in Starch and Elliot's (1912,1913) study. In reference to Brimi's (2011) work, Guskey (2013) made the following comment: "So even if one accepts the idea that there are 100 discernible levels of student writing performance, it's clear that even well-trained teachers cannot distinguish among those different levels with much accuracy or consistency" (p. 70).

The study also shows that grading continues to be subjective and unreliable and that an A in one teacher's class might not be an A in another class or school. Brimi (2011) acknowledged that grades still play an important role in education. Brimi pointed out that grades serve as evidence of student achievement and also help to determine acceptance in many colleges. In addition, Brimi noted that one major change associated with grades today is the "magnitude of the consequences connected to grades" (p. 1). For instance, many scholarships today are closely linked to student grades. Brimi added that as the competitive nature of scholarships increases, the meaning of a grade is called into question more frequently. In response to questions about grades, some states have enacted laws to make grading scales uniform in an attempt to help ensure that the same percentage represents the same letter grade.

As Brimi (2011) explained, the subjectivity associated with grading has various implications. College admission can be linked directly to grades, as can scholarships. If students qualify for acceptance in a college or qualify for a scholarship based on inflated grades, their college experience will most likely be unproductive and could possibly result in students failing to earn their degree. Vatterott (2015) also explained that grades can therefore be misleading when used as a means of predicting success, since inflated grades might not necessarily lead to success in college. This is further supported by Goodwin (2011) who reported that there is a 30% dropout rate among freshmen at 4-year institutions; only 54% of students entering colleges are predicted to complete their degree (Stewart, 2012, as cited in Vatterott, 2015).

The Case Against Percentage Grading Scales

As Guskey (2013) clearly explained, one part of our current grading system which stands as an impediment to making grades more accurate and meaningful is percentage grading (p. 68). Percentage grading, as defined by Brookhart (2004), means assigning grades as percentages which are usually the percent of correct responses on exams or the percent of points earned for assignments (p. 23). Guskey (2013) outlined and explained the four major problems with percentage grading scales. The first problem, according to Guskey (2013), “from the perspective of simple logic is that percentage grading makes little sense” (p. 70). Most teachers who use percentage grades usually set the minimum passing grade at approximately 60 or higher. This implies that there are “60 or more distinct levels of failure and only 40 levels of success” (Guskey, 2013, p. 70). Guskey (2013) went on further to explain that distinguishing among 60 different levels of failure is not very helpful since students who are unsuccessful are usually not concerned with

which level of failure they have achieved. Since no one uses these 60 different levels of failure, Guskey (2013) proposed replacing this 100-point system with a 50-point grading scale having 10 designated levels of failure instead, with the other 40 points describing levels of success/passing and ways students can be successful. With the current percentage grading system and its large number of designated levels of failure, it implies that levels of failure are more distinguishable than the more significant levels of success. Educators, Guskey (2013) argued, who value student learning should be more concerned with distinguishing and describing ways students can achieve success rather than focusing heavily on levels of failure.

The second point Guskey (2013) offered as to the problem of percentage grading is its inaccuracy. It is widely known in science that the accuracy of a measure is dependent on the precision of the measurement instrument. Measurement experts identify precision using the standard error of measurement, which is the amount by which a measure might vary from one occasion to the next using the same device to measure the same trait. Guskey (2013) attested that a percentage grading scale with its 100 classification levels offers the illusion of precision, especially when compared with scales having fewer levels such as excellent, average, and poor. Greater classification levels, Guskey (2013) explained, increase the likelihood of students being misclassified in terms of their performance on an assessment. As an example, a student is more likely to be misclassified as performing at a level of 85% when the true classification is 90% than he is of being misclassified as average when his true classification is excellent. “The increased precision of percentage grades is truly far more imaginary than real” (Guskey, 2013, p. 70).

The third problem Guskey (2013) proposed with percentage grades is that they are misleading. Guskey (2013) explained that a student's percentage grade is usually perceived as a reflection of the percentage of the content the student has mastered. This view is not always true since assessments vary greatly in their design. Some assessments can be very challenging even for students who have achieved mastery of the skill being assessed. This variation in the assessment design can lead to percentage grades which might be misleading.

Finally, Guskey (2004, 2013) highlighted the zero grade as a major problem with percentage grading. Many states, Guskey (2013) explained, have in recent years passed legislations stipulating the minimum grade a teacher may assign, usually 50 on a 100-point scale. The reason for this is inherent in the percentage grading system which describes 60 different levels of failure compared to 40 levels of success. With such a grading system, only a 10-point gain is needed for a student to improve from a B to an A; however, in order to move from the lowest end of the scale, a zero, to the minimum passing grade, a student would need at least six times that amount to move from a B to an A or from a C to a B. Assigning a grade of zero therefore serves as punishment, recovery from which is highly unlikely.

Guskey (2013) proposed a simple solution to solve the problem: simply replace the current 100-point system with a grading system of 0-4. This, he claimed, would make grades more accurate and honest and would also allow students easier recovery from failure.

Reeves (2011), on the other hand, asserted that there is nothing wrong with letter grades. Instead, Reeves believed the toxicity commonly associated with letter grades is

due to the means by which they are reported in the absence of other meaningful and descriptive information.

The main message from many of the assessment experts is that the overall grade or single letter grade is not an ideal way to report on student progress (Brookhart, 2011; Guskey, 2009; Marzano & Hefflebower, 2011). This sentiment has also been echoed by Marzano (2006) who noted that “isolated overall grades (or overall percentage scores)” (p. 125), usually given at the end of a marking period, “are extremely deficient because they cannot provide the level of detailed feedback necessary to enhance student learning” (Marzano, 2006, p. 125).

Notwithstanding all the criticisms over the years, Olson (1995) pointed out the use of grades continues to be “one of the most sacred traditions in American education” (p. 24). Guskey (2015) also explained that this is, in part, due to the many grave consequences and decisions associated with grades. Grades determine if students are promoted to the next grade level. Grades determine student enrollment in advanced classes. Grades also serve as indicators of various disabilities (Guskey, 2015). Today, grading, reporting, and communicating student learning continue to be some of the most seminal responsibilities of educators (Guskey, 1996; Trumbull & Farr, 2000).

Many researchers and grading experts, including Brookhart (2015); Bowers, Sprott, and Taff (2013); Reeves (2008); Guskey (2015); Vatterott (2015); and Dueck (2014), over the years have added to the grading debates and continue to present several salient points which help to validate the need for further studies about grading practices. These experts agree that grades are not only ubiquitous, but they are also major predictors of college acceptance, college performance, college graduation, and also dropping out of

high school. In one recent study, Bowers et al. conducted a wide review of literature and an analysis of 110 dropout indicators across 36 different studies in an attempt to determine the most accurate indicators of students at risk of dropping out of high school. From this study, Bowers et al. concluded that “low and failing grades ... are some of the most accurate indicators of students at risk of dropping out” (p. 97); yet according to Reeves (2008), neither the weight of scholarships nor the many other grave consequences associated with grades seem to have influenced grading policies and practices sufficiently (p. 85). Reeves (2008) further suggested that “if you want to make just one change that would immediately reduce student failure rates” (p. 85) and by extension reduce high school dropout rates, “then the most effective place to start would be challenging prevailing grading practices” (Reeves, 2008, p. 85).

Evidently, the current problems relating to grading and communicating student learning are not new but instead have become chronic issues which have perplexed educators for years (Guskey, 1996, p. 1). Guskey (2015) posited that one of the reasons grading practices have not changed much is due to the “seriousness of the consequences attached to grades” (p. 3) and a fear that changes to these traditional practices might interfere with these consequences. In addition, Vatterott (2015) also acknowledged that reforming grading practices is a difficult challenge because our practices are so deeply rooted in our culture. She went on to conclude that in order for grading reform to occur, educators and all related stakeholders must “acknowledge and accept how our beliefs have influenced grading practices” (Vatterott, 2015, p. 6); however, others like Kohn (2011) believed that the debate and research should not be focused on grading reform or improving grading practices but instead forcefully argued that rather than seeking ways to

improve grading, we should be more concerned about seeking ways to abolish grades altogether. Kohn (2011) accepted that it is important to assess the quality of learning but explained that this does not imply that we need to measure it by converting these measures to numbers. Kohn (2011) further argued that this act of grading, among other vices, compromises student thinking. Kohn's (2011) stance on grade abolition is not unique and has been shared by others like McElligott and Brookhart (2009, as cited in Guskey, 2009), who expressed, "In a perfect world there would be no grades" (p. 67); instead, "students and teachers would work together until students have reached a satisfactory level of achievement of intentional knowledge and skills" (p. 67). However, even those researchers who share this view will also acknowledge that such a statement is utopian in nature and therefore argue for grading reform.

Today many others, including Guskey (2011), continue to reiterate that despite several efforts and the plethora of programs to improve education over the last few decades, grading and reporting is still the one area which remains in serious need of reform. Vatterott (2015) acknowledged that K-12 educational reform is not the solution for all the shortcomings of education in the United States but espoused that the reform of one major educational practice, grading, has the potential to drive related reform in other practices (p. 5).

In addition, Muñoz and Guskey (2015) asserted that when assessments are graded and reported in the correct manner, they can be a "powerful tool for student learning" (p. 67). The effect of feedback and classroom assessments on student learning is a subject that has been widely researched and reviewed by many experts in the field of education. As Marzano (2006) described it, classroom assessments "might be one of the most

powerful tools in a teacher's arsenal" (p. 2). This idea has been widely supported by other assessment experts (Black & Wiliam, 1998; Hattie, 1999, 2009; Stiggins, 2005); however, the common and recurring theme in the literature relating to grading practices is the discrepancy between actual practices and those recommended by grading experts (Brookhart, 1991, 1993, 1994; Guskey, 1994, 2006, 2009; O'Connor, 2011; Vatterott, 2015).

Purpose of Grading

As Guskey (1996, 2000) outlined, there are many purposes for grading and reporting. Grades serve as a means of communicating a student's level of achievement; grades serve as a means of evaluating the effectiveness of programs, providing information to students for self-evaluation; and grades also serve as a means of motivating, sorting, and selecting students for various programs. Guskey (1996) believed that having so many purposes for grading poses a major problem since it results in some teachers emphasizing different purposes compared to other teachers. As a consequence, this causes teachers to use varied criteria to determine a student's grade. This also means that students might receive different grades for the same level of academic achievement (Guskey, 1996).

However, among all the different reasons we grade, the grading purpose that garners the most support from researchers and measurement specialists is student feedback (Troxell, 2009). In addition, Vatterott (2015) also explained that in any standards-based classroom, feedback, not grading, is what matters the most (p. 79). Guskey (1994) supported the idea that grades are not really necessary for learning. This conclusion was also corroborated by Vatterott who explained that what is really

necessary for learning is feedback.

Reeves (2011) also asserted that grades are not inherently toxic. Instead, he explained that it is the absence of additional information which renders the grade inaccurate and misleading (Reeves, 2011, p. 78). Teachers in some of the best classrooms use grades as just one of the many types of feedback given to students. These teachers act on the belief that mistakes made along the way during the grading period are not failures to be averaged into the final grade (Reeves, 2008). When teachers acknowledge that mistakes are an inevitable part of learning, they approach grading with caution while students are still learning. In addition, these teachers provide appropriate, nonpunitive feedback rather than grading students' first attempts on a particular task. This idea is also supported by Kohn (1994) who advised that we should "Never grade students while they are still learning" (p. 41). When teachers use feedback rather than just grades throughout the learning process, students persevere to succeed (Vatterott, 2015). Furthermore, if teachers want students to accept and embrace the idea that mistakes help to facilitate learning, the threat of grading needs to be removed while students are still learning. "We need to teach and grade in ways that garner hope for students, otherwise, they will throw down the ball and go home" (Wormeli, 2006a, p. 18). In a traditional classroom, it is customary to grade students' first attempts on tasks/assessments; but in the standards-based classroom, students are provided with "feedback that is informative and non-judgemental" (Vatterott, 2015, p. 52). The most important activity in the standards-based classroom is not grading but feedback (Vatterott, 2015, p. 79).

Feedback and Standards-Based Grading

Education professor and researcher, John Hattie, is especially known for his extensive work on the various factors that influence student achievement and, in particular, the power of feedback. Hattie (1999, 2009, 2012) noted that over the years, our knowledge in the field of education has grown significantly; and as a result, much is known today about what makes a difference in the classroom. However, Hattie (2009) asserted that despite this rich research base in education today, much of what is known is seldom used by teachers, and many teachers frequently make wild claims about various innovations that seem to work in their classrooms.

According to Hattie (2009), an innovation is described as being “a constant and deliberate attempt to improve the quality of learning” (p. 12). Hattie (2009) acknowledged that it is common in education to hear teachers frequently making statements promulgating claims that their innovation, compared to others, works and enhances student achievement in their class. Hattie (2009) believed that in education, these claims tend to be frequently made because most innovations do indeed have positive influences on student achievement. In an attempt to solve this problem of finding evidence to determine the impact of an innovation on student achievement, Hattie (2009) embarked upon a project to develop a method that would allow for the various innovations to be ranked and therefore compared. In addition, Hattie (2009) argued that the more salient question educators should consider is, “to what extent does an innovation work in the classroom?” Hattie’s (1999, 2012) goal was to develop a way of using over 30 years of educational research to assess the effects of innovations and schooling. This project helped to quantify and rank innovations based on their effectiveness.

In order to determine the impact of schooling on student achievement, Hattie (1999) first needed to establish a benchmark to represent the typical effect of schooling or an innovation on student achievement. The first stage was to establish a continuum summarizing the effects of schooling, ranging from negative values through zero to some positive values. The scale on this continuum would be in terms of effect sizes. A “0” on this continuum would indicate there was no effect of the innovation on student achievement; a positive effect would indicate an increase in student achievement; and a negative effect on this continuum would indicate that the innovation has a decreased effect on achievement. The overall effect sizes related to a particular innovation or influence on student learning were obtained by statistically synthesizing the effects from many studies on a particular topic.

An effect size is described by Hattie (1999) as the “magnitude of study outcomes for all types of outcome variables” (p. 3). Additionally, an effect size of 1.0 is considered to be large and would imply that 95% of outcomes in a study were positively enhanced (Hattie, 1999, p. 3). Hattie (1999) was able to synthesize several studies in order to arrive at an overall conclusion and magnitude of effect. In so doing, he found that there are very few innovations in education that yield an overall negative effect. Hattie (1999) explained that this is one reason why many teachers can boldly state that many innovations they use work for them in their classroom. For instance, a synthesis of meta-analyses on the effects of computers on student achievement revealed that the average effect size across 557 studies was 0.31. This effect size is considered to be so small that it might not be noticeable (Hattie, 1999).

In reference to this, Hattie (2012) made the following statement: “Almost

everyone can impact on learning if the benchmark is set at $d > 0$ " (p. 14). Hattie (1999) therefore concluded that when comparing educational innovations, the reference point should not be at zero but rather a point or effect size which represents the typical effect of schooling. This reference point was taken to be 0.40. This effect size was not arbitrarily determined, but instead was found by averaging the effects across 357 meta-analyses. Hattie (1999, 2009) therefore argued that 0.40 should represent the standard from which all other innovations can be judged. Table 1 shows how this average effect size (ES) was obtained.

Table 1 summarizes Hattie's (2009) synthesis of over 900 meta-analyses which focused on several influences on learning such as the home, school, teacher, and curriculum. This synthesis was also based on more than 50,000 individual studies involving more than 200 million students with ages ranging from 4 to 20 across all subjects.

Table 1

Average Effect for Each of the Major Contributors to Learning

Across dimensions	No. of meta-analyses	No. of Studies	No. of People	No. of effects	ES	SE
Student	152	11,909	9,397,859	40,197	0.39	0.044
Home	40	2,346	12,066,705	6,031	0.31	0.053
School	115	4,688	4,613,129	15,536	0.23	0.072
Teacher	41	2,452	2,407,526	6,014	0.47	0.054
Curricula	153	10,129	7,555,134	32,367	0.45	0.075
Teaching	412	28,642	52,611,720	59,909	0.43	0.070
Average	913	60,167	88,652,074	160,054	0.40	0.061

Source: Hattie (2012, p. 14).

This extensive synthesis revealed that a common feature present in the top influences on student achievement was feedback. Furthermore, several years ago when Hattie (1999) first started to examine the influences on student learning, he made the following statement which since has been quoted widely in research articles and books:

The most single moderator that enhances student achievement is feedback. The simplest prescription for improving education must be “dollops of feedback” – providing information how and why the child understands and misunderstands and what directions the student must take to improve. (p. 9)

Since making these statements endorsing the power of feedback, many more researchers began to study the effects of feedback. After drawing these conclusions about feedback and establishing a scale on the effects of innovations on student achievement, Hattie (2009) has continued to explore the topic of feedback and factors affecting its effectiveness. Years later, after continuing his studies on feedback, Hattie

(2012) expressed regret in making his “dollops” claim relating to the effects of feedback as he believed his statement distracted readers from the more important finding that even though feedback has an overall positive effect, it varies in its influence. Hattie (2012) has since continued to study this variance in an effort to better understand how to improve the effectiveness of feedback.

How Effective is Feedback?

Table 2 shows the effect sizes from 12 meta-analyses assessing the influences of student feedback. The average effect size from studies relating to feedback was 0.79, almost twice the average effect of schooling. This strongly supports the argument that feedback is a powerful influence on student learning. A closer examination of the table also shows great variations in the effect sizes associated with the types of feedback. This illustrates that not all types of feedback are effective. The studies that produced the largest effect sizes were ones in which students received feedback on how to improve (Hattie & Timperley, 2007; Wiliam, 2011, 2016).

Table 2

Summary of Effect Sizes Relating to Types of Feedback

Variable	Number of meta-analyses	Number of studies	Number of effects	Effect sizes
Cues	3	89	129	1.10
Feedback	74	4,157	5,755	0.95
Reinforcement	1	19	19	0.94
Video or Audio feedback	1	91	715	0.64
Computer-assisted instructional feedback	4	161	129	0.52
Goals and Feedback	8	640	121	0.46
Student evaluation feedback	3	100	61	0.42
Corrective feedback	25	1,149	1,040	0.37
Delayed versus immediate	5	178	83	0.34
Reward	3	223	508	0.31
Immediate versus delayed	8	398	167	0.24
Punishment	1	89	210	0.20
Praise	11	388	4,410	0.14
Programmed Instruction	1	40	23	-0.04

Source: Hattie and Timperley (2007, p. 84).

Over the years, it has been shown that feedback is among the most powerful influences on student learning. It has also been shown that the impact is not always positive (Hattie & Timperley, 2007). Additionally, Wiliam (2011, 2016) explained that the process of providing feedback is not as simple as it may seem but further pointed out that “the only important thing about feedback is what students do with it” (p. 10). According to Wiliam (2011), “If I had to reduce all the research on feedback into one simple overarching idea ... it would be this: feedback should cause thinking” (p. 127). Wiliam (2016) went on further to explain that “the main purpose of feedback is to improve the student’s ability to perform tasks he/she has not yet attempted” (p. 12). In order for students to be able to improve their performance and use the feedback provided, teachers should communicate to students the criteria for success. This idea of communicating the purpose of learning a particular task with students is also supported by Vatterott (2015) who expressed these as learning targets and stated that this is a key component in any standards-based classroom. Vatterott explained that when teachers

convert standards into learning targets sometimes expressed as *I can* statements, students understand more clearly what the goal of learning is and what they must be able to do in order to master a standard.

Standards-Based Grading

Most conventional grading practices today involve the use of one grade to sum up achievement in a subject, and that one grade often includes effort, growth, behavior, ability, and other nonachievement factors; however, with the highly recommended, standards-based, learning-focused grading practices, a grade sums up achievement on a specific standard with effort and behavior reported separately (Brookhart, 2011). In addition, Vatterott (2015) explained that standards-based grading is more than just grading, in that it involves a “complete overhaul of the entire teaching-learning process” (p. 26). It is a shift from an obsession about grades at any cost and at the mercy of learning to a culture where students are motivated and driven by learning as defined by performance on standards. Vatterott explained standards-based grading within the context of a standards-based classroom in a manner which suggests that one cannot take place without the other. In other words, Vatterott emphasized that a cultural shift is required in order for this to be realized.

Learning in the Standards-Based Classroom

According to Vatterott (2015), standards-based learning is necessary for the implementation of standards-based grading. The process, Vatterott (2015) explained, begins with standards which express what students should accomplish. The standards are then unpacked to create learning targets that are more accessible to students, so they know what they should learn and be able to do. Ideally, teachers will select tasks which

address a variety of targets in an authentic and logical way which helps facilitate student conceptual learning. In this way, the standards dictate the learning targets which in turn drive the learning tasks (p. 42).

Learning Targets and Feedback

The literature in support of the use of learning targets is rich. In the standards-based classroom, the learning targets are critical since they not only determine the tasks that students are exposed to, but they also dictate the assessments. Hattie (2009) showed that in a synthesis of four meta-analyses consisting of 304 studies, 42,618 people, and 597 effects related to direct instruction, the overall average effect size was 0.59, which was ranked as high. One of the first major steps in this high yielding strategy was that the teacher determines the learning intentions or learning targets for the lesson and makes them transparent to the students (Hattie, 2009, p. 205). Feedback has been shown to be a powerful influence on student achievement (Hattie, 2012; Wiliam, 2016); however, in order for feedback to be effective, it should always relate to the question “Where am I going?” Feedback should always be provided in reference to the learning target or learning goal; hence, specifying and communicating clear learning targets not only helps to determine tasks and assessments but also helps to facilitate effective feedback from and to students (Hattie, 2009, p. 177).

Brookhart (2011), like many others, believed that grades should not be something students earn but instead should be a function of what students learn. Most traditional grading practices today involve the use of one grade to sum up achievement in a subject; and that one grade often includes effort, growth, behavior, ability, and other nonachievement factors as educators typically combine scores from exams, quizzes,

projects, punctuality, homework, class participation, and other individual assessments. Muñoz and Guskey (2015) explained that this combination usually results in a grade that is less meaningful since it is difficult to interpret; however, with standards-based, learning-focused grading practices, a grade sums up achievement on standards with effort and behavior reported separately (Brookhart, 2011). In the standards-based classroom, learning targets are extremely important since they form the bases for determining the assessment and, by extension, grades. In the standards-based classroom, there is a shift from grading assignments to assessing student progress toward the learning targets. Unlike traditional classrooms, in a standards-based classroom, a student's final grade is determined by the extent to which the student reaches those targets (Vatterott, 2015, p. 46).

According to Muñoz and Guskey (2015), the first step in ensuring sound grading practices is to make them more meaningful. Muñoz and Guskey contended that in order to make grades more meaningful, separate grades should be assigned for achievement, effort, and progress. Muñoz and Guskey suggested that when assigning grades, educators should distinguish among product, process, and progress learning criteria. Muñoz and Guskey further explained each of these in the following ways.

Product criteria. Educators who believe that the primary purpose of grading is to communicate student summative overall achievements on specific goals will assign grades as a reflection of the final product of a student's work. Emphasis is placed on final exam grades, projects, and other forms of summative assessments.

Process criteria. There are many teachers who believe that the process of how the student determines the final product is also important and should be considered.

These educators would prefer to include grades based on the process, and this is referred to as process criteria. This involves grades based on responsibility, effort, work habits, homework, class participation, classroom quizzes, and various other types of formative assessments.

Progress criteria. In addition, there are other educators who believe that the most important message grades should convey is the amount of progress a student has made over a period of time rather than their final grade or how they got there. These educators base their grades on progress criteria. This is also known as value-added learning, learning gain, improvement scoring, or educational growth.

As Muñoz and Guskey (2015) explained, the most important aspect of this approach to grading is that teachers report grades for each criterion separately, thereby ensuring that the grades are more meaningful, reliable, and fair. Another benefit is that there will be a stronger correlation between product grades and state assessments, assuming classroom assessments are aligned to state standards.

Most states today have common standards that address what students should learn/be able to do (Muñoz & Guskey, 2015). With standards-based grading, there are multiple achievement grades reported per subject, since a grade represents a student's achievement on a specific standard. Other grades based on process criteria like effort and behavior are reported separately with standards-based grading. As Brookhart (2011) explained, in order for districts to successfully implement standards-based grading, they must make the shift from traditional grading practices, in which one final grade is reported for a subject which typically includes achievement and nonachievement grades combined to standards-based grading. In addition, this shift from traditional grading

practices to standards-based grading would also imply that districts would have to make adjustments to the way report cards are designed to enable educators to report multiple grades for different learning criteria. As Muñoz and Guskey (2015) reported, very few states have made this shift in reporting grades.

Vatterott (2015) offered hope for district leaders or individual teachers who are interested in making the shift from traditional grading practices to standards-based grading. Vatterott explained that the transition does not have to be grandiose but rather can take place gradually. Vatterott offered advice based on lessons from schools which have implemented standards-based grading.

1. Schools can start small with a grade level, team, or even an individual teacher piloting this reform. Eventually, improvements can be made before expanding to other grades and the rest of the district.
2. Leaders who have successfully implemented standards-based grading have done so with the understanding that it will take time to change a practice that has been so deeply rooted in tradition. These leaders have given teachers time to observe and learn more about the benefits of such a grading system before making any changes mandatory.
3. The research on standards-based grading should not be confined to special committees but should be shared among all stakeholders.
4. As grading expert Brookhart (2011) explained, grading conversations can become heated if they are not conducted in the correct manner. Brookhart advised that these conversations should begin with questions to help establish what teacher beliefs are about the meaning grades should convey. Eventually,

all stakeholders should work together to establish a shared vision or agreed upon principles outlining what everyone believes about grades (Vatterott, 2015, p. 94)

5. Communicate effectively to all stakeholders about the research supporting grading reform via a district-level representative, messages to parents, and also school/district websites.

Scriffiny (2008) explained that if a district or school has not formally embraced standards-based grading, teachers can still make small changes to their grade book that would allow them to report student grades based on standards. Table 3 compares the traditional style of reporting grades to standards-based reporting. This method helps to eliminate the inappropriate use of averages.

Table 3

A Comparison of Traditional and Standards-Based Grade Books

Traditional Grade Book			
Name	Homework Average	Quiz 1	Chapter 1 test
John	90	65	70
Bill	50	75	78
Susan	110	60	62
Felicia	10	90	85
Amanda	95	100	90

Standards-Based Grade Book			
Name	Objective 1: Write an alternative ending for a story	Objective 2: Identify the elements of a story	Objective 3: Compare and contrast two stories
John	Partially Proficient	Proficient	Partially Proficient
Bill	Proficient	Proficient	Partially Proficient
Susan	Partially Proficient	Partially Proficient	Partially Proficient
Felicia	Advanced	Proficient	Proficient
Amanda	Partially Proficient	Advanced	Proficient

Source: Scriffiny (2008, p. 72).

Many grading experts agree that improving grading practices has the potential to reduce failure rates and improve discipline, student motivation, and work ethic and increase faculty morale (Guskey, 2015; Marzano, 2006; Reeves, 2012; Vatterott, 2015).

Thiele (2016) explained the philosophical association between standards-based grading and constructivism:

Lying in the constructivist approach, learning is formed in many ways and is not a permanent quality of a person; standards-based grading practices allow students to see their learning and to work with those around them to continue learning.

Traditional grading practices primarily align with the fixed mindset, by representing each attempt at learning as a final, fixed, mark in the grade book. (p. 8)

Fixes for Broken Grades

O'Connor (2011) offered 15 fixes for grades, as listed below:

Fixes for Practices That Distort Achievement:

- Fix 1: Don't include student behaviors (effort, participation, adherence to class rules, etc.) in grades; include only achievement.
- Fix 2: Don't reduce marks on "work" submitted late; provide support for the learner.
- Fix 3: Don't give points for extra credit or use bonus points; seek only evidence that more work has resulted in a higher level of achievement.
- Fix 4: Don't punish academic dishonesty with reduced grades; apply other consequences and reassess to determine actual level of achievement.
- Fix 5: Don't consider attendance in grade determination; report absences

separately.

- Fix 6: Don't include group scores in grades; use only individual achievement evidence.

Fixes for Low-Quality or Poorly Organized Evidence

- Fix 7: Don't organize information in grading records by assessment methods or simply summarize into a single grade; organize and report evidence by standards/learning goals.
- Fix 8: Don't assign grades using inappropriate or unclear performance standards; provide clear descriptions of achievement expectations.
- Fix 9: Don't assign grades based on a student's achievement compared to other students; compare each student's performance to preset standards.
- Fix 10: Don't rely on evidence gathered using assessments that fail to meet standards of quality; rely only on quality assessments.

Fixes for Inappropriate Grade Calculation

- Fix 11: Don't rely only on the mean; consider other measures of central tendency and use professional judgment.
- Fix 12: Don't include zeros in grade determination when evidence is missing or as punishment; use alternatives, such as reassessing to determine real achievement, or use "I" for Incomplete or Insufficient Evidence.

Fixes to Support Learning

- Fix 13: Don't use information from formative assessments and practices to determine grades; use only summative evidence.
- Fix 14: Don't summarize evidence accumulated over time when learning is

developmental and will grow with time and repeated opportunities; in those instances, emphasize more recent achievement.

- Fix 15: Don't leave students out of the grading process. Involve students; they can—and should—play key roles in assessment and grading that promote achievement. (p. 13)

Why Grading Reform Frequently Fails (Lessons Learned)

Vatterott (2015) explained that the standards-based grading movement is taking place with greater frequency all across the country but cautioned that it has not been without some casualties. Some of the causes for failure are outlined below.

Failure to establish a shared vision and belief statements. Brookhart (2011) emphasized that in order for grading reform to be successful, it is imperative that districts begin by having meaningful and honest conversations about what grades should mean. She explained that in order for these conversations to be productive, they must address long-held beliefs and deeply entrenched practices of teachers respectfully. Brookhart (2011) explained that the details about what should count or be factored in a grade should be secondary and can be guided by the agreed upon belief statements.

This conclusion is also supported by Vatterott (2015) who continuously reminded us that “for grading reform to happen, we must acknowledge and accept how our beliefs have influenced our practices” (p. 6). This is also similarly expressed by Guskey (2015) in the following statement:

Reform initiatives that set out to improve grading and reporting procedures *must* first begin with comprehensive discussion about the purpose of grades and of the report card. These discussions should focus on the message to be communicated

through grading and reporting, the audience or audiences for that message, and the intended goal of the communication. Once decisions about purpose are made, other critical issues about the form and structure of the report card, as well as issues related to broader grading and reporting policies and practices, will be much easier to address and resolve. (p. 21)

Implementing changes too quickly. “Many school districts that have struggled to implement standards-based grading have suffered from an approach that was ... too fast, too much at once” (Vatterott, 2015, p. 102). Furthermore, Vatterott (2015) asserted that when teachers and other stakeholders begin to embrace standards-based grading, the process involves so much more than just grading. Grading reform efforts have the potential to drive major changes in other practices related to instruction. Similarly, Brookhart (2011) also explained that when districts make the decision to move away from traditional grading practices, their next steps usually involve professional development not just about how to implement aspects of the new grading policy but also professional development in teaching and learning strategies. Brookhart (2011) explained that when districts begin to grapple with and facilitate honest conversations about grading, questions about learning strategies inevitably emerge. This is due to the fact that with standards-based grading, there is a greater emphasis being placed on learning and progress toward mastery of standards. Consequently, when students do not master a standard, teachers begin to seek support in finding different instructional pathways to assist students meet those standards.

When the implementation of standards-based grading is too fast, teachers might not receive the professional development needed to support the necessary changes. In

addition, when the process is too fast, it might give parents and other stakeholders the impression that there was no pilot program or phase-in stage to help manage the transition. It could also appear that communication was lacking and that the necessary groundwork was poorly laid (Vatterott, 2015).

Summary

This review of literature has revealed several distinct differences between traditional grading and the less common but more desirable alternative, standards-based grading.

The purpose of this qualitative study was to describe teacher current classroom grading practices and to explain the process by which teacher grading practices develop and evolve throughout their teaching careers. In addition, the study explored the reasons why teachers persist in embracing outdated traditional grading practices. The specific research questions that guided this study were

1. What are the grading practices currently used by teachers?
2. What factors influence the development and evolution of teacher grading practices?
3. What changes (if any) might be implemented in current grading practices which would provide a better measure of assessing student achievement?

In one publication by the Southern Regional Education Board (n.d.), 10 strategies are outlined for creating a culture of high classroom expectations. The ninth strategy which focuses on the importance of establishing sound grading practices, represents a fitting summation of this literature review on grading and an appropriate conclusion to this chapter:

One of the most misused, misunderstood and mistrusted issues in public schooling is how we have communicated student achievement and progress to our publics.

Grades must communicate to parents, students and teachers exactly what students know and are able to do. Grades must also communicate what quality work looks like. Therefore, we need to show students what good work is — whether in mathematics, music, English or auto mechanics. They need scoring guides that clearly delineate what the expectations are for earning top grades (As, Bs); acceptable grades (Cs); and Not Yet! (NY) for work that is not accepted.

Finally, the culture of assessment needs to reflect that every student is capable of — and expected to — produce acceptable work. Sub-standard work will not be accepted (Not Yet!) as final until it is at least minimally acceptable. This type of grading embraces the conviction that all students can and will learn. It decreases frustration and shifts the responsibility for grades earned from the teacher's red pen to the student's choices and degree of effort. (Southern Regional Education Board, n.d., p. 7)

Chapter 3: Methodology

Chapter 3 presents the methodology and procedures that were implemented for this study. In addition, this chapter explains and outlines the purpose of the study, participants, the research design, instrumentation, data collection, data analysis, limitations, and delimitations. The chapter concludes with a summary.

Introduction

There is still much that is unknown about grading and reporting and “conclusive evidence to identify precisely what is truly *best* practice” has not yet been established (Guskey, 2015, p. 109). Creswell (2012) explained that a qualitative research design “is best suited when the literature might yield little information about the phenomenon of study, and you need to learn more from participants through exploration” (p. 16). Accordingly, the study was conducted by using a qualitative research design and employed the methods of the grounded theory approach.

According to Guskey (2009, 2015), “of all aspects of our education system, none seems more impervious to change than grading and reporting” (p. 2). The study explored the extent to which this statement reflects the practices of the teacher participants or if, on the contrary, their practices have evolved to include more nontraditional methods. Furthermore, if teacher grading practices included features of nontraditional grading, it would be worthwhile to learn more about the various influences that have led to this change; however, if teacher grading practices have indeed been resistant to change, as the literature suggests, the researcher sought to learn more about the reasons behind any such resistance.

The purpose of this qualitative study was is to describe teacher current classroom

grading practices and explain the process by which teacher grading practices develop and evolve throughout their teaching careers.

Researchers continue to report that traditional approaches to grading and reporting are still commonly practiced in most schools today and represent an enduring problem (Guskey, 2015; Vatterott, 2015). Traditional practices such as averaging to calculate a final score, the use of zero on a 100-point scale, and the inclusion of student behaviors in grade calculation such as giving grade penalties for late work or extra credits for compliance are all considered to be counterproductive to student learning (O'Connor, 2011; Reeves, Jung, & O'Connor, 2017; Schimmer, 2016; Vatterott, 2015).

Standards-based grading has been proposed as the more favorable alternative to address problems associated with traditional grading (Guskey, Swan, & Jung, 2011; Vatterott, 2015). Characteristics of standards-based grading include the assignment of grades based on level of mastery of learning targets not tasks; the practice of not grading homework and other formative assessments; the replacement of old information with more recent grades as a student demonstrates evidence of additional learning; grading student learning only; the provision of multiple opportunities for students to improve through the use of formative assessments; and feedback that is “free from the threat of grades” (Vatterott, 2015, p. 79).

Muñoz and Guskey (2015) explained that grading and reporting remain powerful tools to improve student learning. Similarly, Reeves (2008) purported that when we “challenge prevailing grading practices” (p. 85), there is a corresponding reduction in failure rates among students. These statements from researchers support the claim by Brookhart et al. (2016) that grades continue to play a central role in a student's

educational experience and remain an important topic to be studied.

The three questions which emerged from the review of the literature on grading and which are of great interest to the researcher are

1. What are the grading practices currently used by teachers?
2. What factors influence the development and evolution of teacher grading practices?
3. What changes (if any) might be implemented in current grading practices which would provide a better measure of assessing student achievement?

These research questions served as a guide for the study and were resolved through analysis of the data collected. The research methodology that was used to address the research questions was a qualitative approach which employed the methods of grounded theory.

Research Design

There are several different approaches to qualitative research. Narrative research describes the stories of participants; phenomenology highlights the common experiences of a number of individuals, but the grounded theory approach aims to “move beyond description ... to generate or discover a theory” (Creswell & Poth, 2018, p. 82).

Grounded theory is a form of qualitative methodology which enables the researcher to develop theory about the central phenomenon that is firmly grounded in the data collected (Charmaz, 2014; Corbin & Strauss, 2015; Creswell, 2012). This methodology was developed in the late 1960s by two sociologists, Barney G. Glaser and the late Anselm L. Strauss, and evolved from their work with patients who were terminally ill. Their research methods which were subsequently published attracted the attention of

others who wanted to learn more about their grounded theory procedures. They later published a book detailing their research methods. In their book, Glaser and Strauss (as cited in Creswell, 2012) posited that discovering a theory that is grounded in actual data from participants is far more informative than simply verifying and testing theories that were identified before the start of the study. Glaser and Strauss's book outlines the major ideas of grounded theory practiced today and has served as a guide for many dissertations (Creswell, 2012).

The grounded theory approach to qualitative research has been in use for a long time and consequently offers the added advantage of providing "a tried-and-true set of procedures for constructing theory from data" (Corbin & Strauss, 2015, p. 11).

The purpose of this grounded theory study was to describe teacher current classroom grading practices and explain the process by which teacher grading practices develop and evolve throughout their teaching career. It is the hope of the researcher that through this study more knowledge will be generated to enable stakeholders to have a greater positive influence on teacher grading practices. Additionally, another desired outcome of this study was that the data generated will help to galvanize principals, district leaders, and other policy makers to take action in addressing "antiquated practices" (Guskey, 2009, p. 2). According to Corbin and Strauss (2015), "The knowledge gained through grounded theory methodology enables persons to explain and take action to alter, contain, and change situations" (p. 11). Grounded theory, therefore, was an appropriate methodology for this study.

Corbin and Strauss (2015) contended that grounded theory methods are ideal in providing "new insights into old problems as well as to study new and emerging areas in

need of investigation” (p. 11). Traditional grading practices remain an old and enduring problem in contrast to standards-based grading, the relatively new and emerging alternative. According to Guskey (2015), “There is much about grading and reporting we still don’t know. We don’t yet have sufficiently conclusive evidence to identify precisely what is truly *best* practice. It remains an area ripe for careful study and thoughtful investigation” (p. 109).

It is hoped that through this study new knowledge will be acquired to not only explain the process by which teacher grading practices develop and evolve throughout their teaching careers but also to equip stakeholders with more effective ways of influencing grading practices positively.

Participants

The target population for this study was approximately 30 middle school core-academic teachers from Title I schools in a large school district in North Carolina.

Rationale for Target Population

According to the U.S Department of Education (2018),

Title I, Part A (Title I) of the Elementary and Secondary Education Act, as amended by the Every Student Succeeds Act (ESEA) provides financial assistance to local educational agencies (LEAs) and schools with high numbers or high percentages of children from low-income families to help ensure that all children meet challenging state academic standards. (para. 1)

A review of the literature related to grading suggests that effective grading practices can improve student academic achievement (Marzano, 2000, 2006; Reeves (2011, 2012, 2016; Reeves et al., 2017; Vatterott, 2015) and is therefore an ideal topic to study in Title I schools where the need to focus on school reform efforts that yield greater

student achievement is even more urgent. More recently, Reeves (2016) explained that “when schools embrace effective grading practices and terminate toxic grading policies, student performance, motivation, and discipline improve” (p. 2).

Grading and reporting remain areas overlooked in many school reform efforts (Guskey, 2011; Vatterott, 2015). Furthermore, Title I schoolwide reform programs exist for the purpose of improving achievement of the lowest achieving students. The data from this study could provide useful information pertaining to teacher current grading practices and could help determine if there is a need to focus on grading and reporting as a reform effort in Title I schools in this district.

Information from the district website for the period 2017-2018 revealed the following: The district is the second largest employer in the county with more than 3,700 full-time and part-time employees, including 1,950 classroom teachers. The district has a student population of over 32,000 which includes over 6,000 middle school students. More than 65% of the district’s students receive free or reduced priced lunches. Over 59% of the students in the school district are classified as Caucasian/White; just above 21% are classified as African American/Black; approximately 4% are classified as multiracial; and the remainder of the students in the school district are classified in racial/ethnic categories of Hispanic (>12%), Asian (>1%), and American Indians. Hawaiian or Pacific Islanders comprise approximately 0.3% of the district population.

Purposeful Sampling

Purposeful sampling was the method that was used to select the participants or sites for this qualitative study. According to Creswell (2012), “In qualitative inquiry, the intent is not to generalize to a population, but to develop an in-depth exploration of a

central phenomenon” (p. 206). Furthermore, Creswell made the following distinction between quantitative and qualitative sampling:

In quantitative research, the focus is on random sampling, selecting representative individuals and then generalizing from these individuals to a population.... The research term used for qualitative sampling is purposeful sampling. In purposeful sampling the researcher intentionally selects individuals and sites to learn or understand the central phenomenon. (p. 206)

The central focus of this study was to explain the process by which teacher grading practices develop and evolve throughout their teaching careers. The study described and reported on the current grading practices of middle school core academic (math, language arts, social studies, science) teachers in one school district in North Carolina. In addition, school and district grading policies were examined in order to determine the extent to which these policies and practices aligned to those recommended by research and the extent to which these and other factors influenced the development of teacher grading practices.

Middle school core academic teachers who taught mathematics, English, science, or social studies were targeted for this study, in contrast to one subject area, because this provided more informative data that allowed for comparisons to be made among the different subject areas. Since the central phenomenon was the development and evolution of middle school teacher grading practices, participants were purposefully selected to ensure that their teaching experience, in years, was dissimilar.

Sampling middle school teachers was appropriate for this qualitative study since they provided information that explained their grading practices. In addition,

interviewing administrators for these schools also provided useful information about grading policies at the school and district level that were contributing factors in the development of teacher grading practices.

When permission to conduct the study was sought and granted by the IRB and the district, middle school core academic teachers who taught mathematics, English, science, or social studies were invited to participate in the study via email. Middle school teachers with varying years of experience were purposefully selected to participate in the study. A sample teacher consent form can be found in Appendix A.

Theoretical Sampling

In narrative studies, the researcher purposefully selects individuals who can offer stories about their experiences. In phenomenological studies, the researcher purposefully selects participants who have experienced the central phenomenon (Creswell & Poth, 2018); however, in grounded theory studies, a unique form of purposeful sampling known as theoretical sampling is also used (Corbin & Strauss, 2015). Theoretical sampling is a sampling strategy that is specific to grounded theory and which allows the researcher to intentionally and flexibly seek and collect data that is of theoretical value (Charmaz, 2014).

In this study, purposeful sampling was used in the initial stages of the study, but as some data was collected, theoretical sampling strategies was employed to further develop the initial concepts identified.

Theoretical sampling was further explained by Charmaz (2014) who cautioned against mistaking theoretical sampling with other forms of sampling such as sampling to address initial research questions or sampling to reflect population. These, Charmaz

explained, are the initial types of sampling: “Initial sampling provides a point of departure, not of theoretical elaboration and refinement. We cannot assume to know our categories in advance.... Initial sampling in grounded theory gets you started; theoretical sampling guides you where to go” (p. 197).

In the initial stages of this study, Title I middle school core academic teachers of math, English, science, and social studies were purposefully selected to participate in the study. After the first set of data were collected, categories of information emerged from the data analysis. The researcher gathered more data that helped to develop these initial concepts/categories (Corbin & Strauss, 2015). In order for these preliminary categories to have developed into more “robust categories that stand on firm, not shaky, ground” (Charmaz, 2014, p. 192), theoretical sampling was used later in this grounded theory study to ensure that all data that were subsequently gathered were focused on developing those categories and generating a theory.

Instrumentation

In this study the sources of data included but were not limited to one-on-one interviews, focus group interviews, questionnaires, surveys, documents, and research journals/diaries.

Creswell and Poth (2018) explained the conventional use of instruments in qualitative studies:

The qualitative researcher collects data themselves through examining documents, observing behavior, and interviewing participants. They may use an instrument but it is one developed by the researcher using open-ended questions. They do not tend to use or rely on questionnaires or instruments developed by other

researchers. (p. 43)

The primary data collection instrument for this study consisted of one-on-one interviews with teachers and principals. Since “qualitative research is an emerging design” (Creswell, 2012, p. 130), the qualitative data sources included semi-structured interviews with teachers in the district. This also implies that the procedures and questions were not rigid, but instead, could change once the researcher entered the field and started the data collection process. The interviews were for the purpose of obtaining more in-depth information about participant grading practices. An interview guide was used during the interview (see Appendix B for a sample teacher interview protocol).

Pilot test. A pilot test was conducted on the qualitative interview questions to identify and measure the interview content based on the specifications of the study. Three middle school teachers were asked to review and make anecdotal comments about the (a) clarity of the interview questions, (b) bias of the interview questions, and (c) overall flow of the interview questions based on the topic of teacher grading practices. The duration of the interview was noted. Responses from the pilot program were used to revise the interview protocol (shown in Appendix B) to ensure that the clarity/intent of the questions matched the stated purpose of the instrument.

A secondary focus of this study was to examine district and school level grading policies to determine the degree to which those policies influenced teacher grading practices. In addition, the researcher conducted interviews with principals and district leaders to further address the research questions. Sample principal and district leader interview protocols are located in Appendices C and D respectively.

Creswell (2012) explained that data collection in a qualitative research study

“relies on general interviews or observations so that we do not restrict the views of the participants” (p. 205). Creswell further explained that the data collection instrument is researcher designed with open-ended questions.

According to Charmaz (2014), “The question of how many interviews a researcher should conduct pervades qualitative research and remains contested among grounded theorists” (p. 105). Grounded theorists sometimes embrace contradicting positions related to the number of interviews needed. To illustrate, Charmaz explained that some grounded theorists focus on saturating concepts and therefore “eschew attention to large interview samples” (p. 107). In contrast, other grounded theorists highlight constant comparison methods and therefore require a greater number of interviews.

Creswell (2012) explained,

It is typical in qualitative research to study a few individuals or a few cases ... because the overall ability of a researcher to provide an in-depth picture diminishes with the addition of each new individual or site. One objective of qualitative research is to present the complexity of a site or of the information provided by individuals. (p. 209)

Furthermore, Charmaz (2014) offered the following guidelines: “Increase your number of interviews when you: pursue a controversial topic; anticipate or discover surprising or provocative findings; construct complex conceptual analyses; use interviewing as your only source of data; and seek professional credibility” (p. 108).

In keeping with these general guidelines approximately 30 participants, including middle school core academic teachers from Title I schools, school administrators, and

district leaders were targeted for this study.

Procedures

The following procedures were used to conduct the study.

Step 1: Ethical considerations. Permission to conduct the study was sought from the Gardner-Webb University's Institutional Review Board and from the school's district central office. The researcher adhered to all ethical guidelines and policies regarding research.

Figure 1 illustrates the data collection activities that were considered throughout this study (Creswell & Poth, 2018, p. 149).

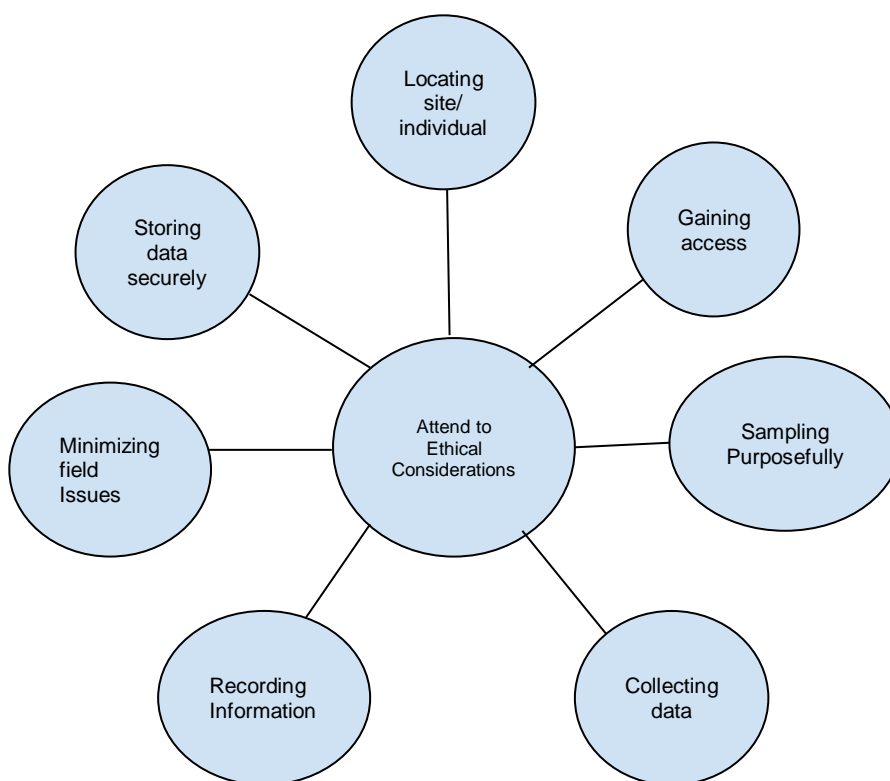


Figure 1. Data Collection Activities. Source: Creswell & Poth (2018, p. 149).

Figure 1 shows that ethical considerations were considered at all stages of data collection and was therefore at the center of the figure. Creswell and Poth (2018)

outlined three main guiding principles to help ensure that the research is conducted in an ethical manner: “respect for persons ... concern for welfare ... and justice” (p. 151). The researcher adhered to these three guiding principles throughout the study.

Informed consent from all participants was sought before data collection began. Measures taken to protect and respect the privacy of participants were clearly communicated and participants were informed of their right to withdraw from the study at any time. Pseudonyms were used to represent participants, places, schools, and school districts associated with the study. A sample of the letter used to seek permission to conduct the study can be found in Appendix E.

Data were collected mainly through interviews which were tape recorded and later transcribed verbatim. The researcher communicated the purpose of the study and sought the consent of participants before data collection began. The researcher also explained to the participants that data would be coded and stored securely in a place where only the researcher had access. The interview protocol (included in Appendix B) outlined the purpose of the study, stated how data were protected, and specified that written consent was required before the start of the interview.

Step 2: Conduct purposeful sampling. When permission to conduct the study was granted, participants and sites were selected by purposeful sampling. Purposeful sampling is a form of data collection that is used in most qualitative studies. In purposeful sampling, the researcher intentionally chooses individuals and sites that will provide information about the central phenomenon being studied (Creswell, 2012; Creswell & Poth, 2018).

Teachers were invited through email to participate in an interview. Corbin and

Strauss (2015) contended that in grounded theory studies, “there is no definite number of participants” (p. 135), because researchers need flexibility to sample and explore initial concepts until saturation is achieved. In addition, Creswell and Poth (2018) explained that in grounded theory studies, the final sample size is not determined by a number but instead is determined when theoretical saturation is reached.

In keeping with these guidelines, approximately 30 middle school teachers, school administrators, and district leaders with varying years of experience were purposefully selected to participate in the study. Principals from the Title I schools and district leaders were also interviewed to support the theory.

Step 3: Data collection and coding. After the first interview was conducted and any analysis of the data was conducted, the entire interview and notes were read and reviewed by the researcher. This enabled the researcher to “enter vicariously into the life of the participants, feel what they are experiencing, and listen to what they are saying through their words or actions” (Corbin & Strauss, 2015, p. 86). Following this initial reading, the data were coded.

Corbin and Strauss (2015) defined coding as the process of “delineating concepts to stand for interpreted meaning of data” (p. 220). After the first data were collected, the process of coding began so the researcher could decide what data needed to be collected next. Creswell (2012) explained, “a characteristic of grounded theory research ... is that the inquirer collects data more than once and keeps returning to the data sources for more information until all categories are saturated and the theory is fully developed” (p. 441).

The researcher analyzed the data for emergent themes and made adjustments to the interview questions when needed. Data collection and data analysis took place

constantly until saturation was reached (Creswell, 2012).

Creswell (2012) explained that in grounded theory studies, a unique form of purposeful sampling known as theoretical sampling is used and which “begins after the first analytical session and continues throughout the entire research process” (Corbin & Strauss, 2015, p. 140). With this form of sampling method, data and concepts are sampled for its theoretical value (Corbin & Strauss, 2015; Creswell, 2012).

Accordingly, after the first set of data were collected and initial categories were identified, theoretical sampling was employed thereafter to further explore and develop those categories. Creswell (2012) explained, “Theoretical sampling in grounded theory means the researcher chooses forms of data collection that will yield text and images useful in generating a theory. This means that the sampling is intentional and focused on the generation of a theory” (p. 433).

Data collection in theoretical sampling is purposefully open and flexible in order to enable the researcher to follow up on leads which could contribute to the developing theory (Corbin & Strauss, 2015). It is important to note that “unlike statistical sampling, theoretical sampling cannot be planned before embarking on a study” (Corbin & Strauss, 2015, p. 147). Furthermore, Creswell and Poth (2018) explained that the “initial plan for research cannot be tightly prescribed and that all phases of the process may change or shift after the researcher enters the field and begins to collect data” (p. 44).

As Corbin and Strauss (2015) explained, “A researcher using theoretical sampling never knows what twists and turns the research will take” (p. 137). This implies that the researcher follows general procedures and remains flexible during

the data collection and analysis phase of the study, thereby allowing concepts to drive the data collection process (Corbin & Strauss, 2015).

Step 4: Data analysis (an emerging design). An emerging design in grounded theory refers to the process of analyzing data immediately after it is collected, as opposed to collecting all data before any analysis can begin. The initial analysis will inform the researcher about what type of data is to be collected next (Creswell, 2012).

Following the methods of an emerging grounded theory approach, all data were analyzed using the constant comparison method, coding, memo-writing, and theoretical sampling.

Data analysis in grounded theory studies is explained by Corbin and Strauss (2015) in the following manner:

Unlike conventional methods of sampling, researchers do not go out and collect all the data before beginning the analysis. Analysis begins after the first data are collected. Data collection is followed by analysis. Analysis leads to concepts. Concepts generate questions. Questions lead to more data collection so that the researcher can learn more about those concepts. This circular process continues until the research reaches the point of saturation – that is, the point in the research when all major categories are fully developed, show variation, and are integrated. (p. 135)

In this grounded theory study, the data analysis process will occur once the first set of data are collected from the first interview and will continue throughout the entire data collection phase using a constant comparative method. Creswell (2012) explained the constant comparison method as follows:

In grounded theory research, the inquirer engages in a process of gathering data, sorting it into categories, collecting additional information, and comparing the new information with emerging categories. This process of slowly developing categories of information is the constant comparison procedure. (p. 434)

Creswell further explained, “Constant comparison is an inductive data analysis procedure in grounded theory research of generating and connecting categories by comparing incidents in the data to other incidents, incidents to categories, and categories to other categories” (p. 434).

Analysis, according to Corbin and Strauss (2015), is “the act of taking data, thinking about it, and denoting concepts to stand for the analyst’s interpretation of the meaning intended by the participant” (p. 85).

Additionally, memos and diagrams served as the main analytical tools used in this research study. Corbin and Strauss (2015) defined memos as “written records of analysis” (p. 106) and diagrams as “conceptual visualizations of data” (p. 123). Further, Charmaz (2014) explained memo-writing in the following way:

Memo-writing constitutes a crucial method in grounded theory because it prompts you to analyze your data and codes early in the research process.... Memos catch your thoughts, capture the comparisons and connections you make, and crystallize questions and directions for you to pursue. (p. 162)

Memos were used to record the researcher’s thinking as the data were being analyzed and coded. Each memo was dated and titled with the concept that represented and explained the researcher’s interpretation of the data. The memos served as a record of the process that the researcher used to code the data and also documented how the researcher’s

thinking developed as more data were collected and analyzed.

Stages of data analysis. In grounded theory studies, data analysis occurs in three stages of coding: open coding, axial coding, and selective coding (Corbin & Strauss, 2015; Creswell & Poth, 2018).

Open coding. After the first interview, the first stage of coding will begin and is known as open coding. During this phase, the grounded theorist will examine data in order to identify initial categories or concepts of information related to the phenomenon being studied. In the open coding phase, subcategories known as properties will also be identified which will help to explain the initial categories/concepts (Creswell, 2012). The concepts identified represent summaries of the data collected from interviews and memos (Corbin & Strauss, 2015). The number of categories suggested by Creswell (2012) is 10. Creswell and Poth (2018) further explained the open coding stage: “Using the constant comparative approach, the researcher attempts to “saturate” the categories – to look for instances that represent the category and to continue looking (and interviewing) until the new information obtained does not provide further insight into the category” (p. 203).

Axial coding. In the second phase of coding, the researcher will select one category which can be related to all other categories. During this phase, a diagram known as the coding paradigm is usually created which shows how the various categories are related to each other (Creswell, 2012). The diagram, also referred to as “a coding paradigm” (Creswell & Poth, 2018, p. 203), represents the theoretical model.

Selective coding. During this third phase of coding, known as the selective coding, the grounded theorist will use the information from the axial coding phase to write a theory explaining the interrelationships among the categories in the coding

paradigm. This theory represents an explanation of the process being studied (Creswell, 2012; Creswell & Poth, 2018).

Step 5: Validate the theory. Throughout the entire study, steps were taken to ensure that the data collected and interpretations made were accurate. A variety of data sources were used in order to facilitate triangulation. Creswell (2012) explained triangulation as “the process of corroborating evidence from different individuals, ... types of data, ... or methods of data collection” (p. 259). Furthermore, “the intent of validation is to have participants ... or the data sources themselves provide evidence of the accuracy of the information in the qualitative report” (Creswell, 2012, p. 262).

Step 6: Write a grounded theory report. The grounded theory report was written in the form of a narrative report explaining the relationships among the categories identified.

Delimitations and Limitations

The following represents delimitations in the study:

1. Only core subject teachers were sampled for this study; that is math, language arts, science, and social studies teachers.
2. Only teachers in Title I middle schools were invited to participate in the study.

This represents four middle schools of a total of 11 in the district.

Participation in this study was voluntary and could possibly affect the sample size. The sample size depended on the number of respondents and a small sample and could represent a limitation for this study.

Charmaz (2014) explained, “grounded theorists try to elicit their interviewees’

stories, to the extent that they are willing to share them” (p. 87). This dependency on participant stories that is inherent in grounded theory methods represents a possible limitation in the study if participants do not share their experiences openly, completely, and honestly.

Summary

The chapter commenced with an introduction which included a summary of the literature on grading. The literature reveals that there is much to be learned about what exactly constitutes best grading practices. In addition, the literature suggests that traditional practices, though less desirable than standards-based grading, are still commonly practiced today. The purpose of this grounded theory study was to describe teacher current classroom grading practices and to explain the process by which teacher grading practices develop and evolve throughout their teaching careers. In addition, the study sought to explore the various influences on teacher grading practices and understand the causes for teacher reluctance to modernize their grading practices.

The chapter also presented the research questions which emerged from the literature and which guided this study. The research questions explored were

1. What are the grading practices currently used by teachers?
2. What factors influence the development and evolution of teacher grading practices?
3. What changes (if any) might be implemented in current grading practices which would provide a better measure of assessing student achievement?

Additionally, Chapter 3 outlined the grounded theory methods in detail and discussed the procedures that were used including the data collections and analyses strategies to resolve

the research questions.

Chapter 4: Results

This study sought to describe teacher grading practices and explain the process by which teacher practices develop and evolve throughout their teaching career. The research methodology used to address the research questions was a qualitative approach which employed the methods of grounded theory. This chapter presents the findings of this study in relation to the research questions:

1. What are the grading practices currently used by teachers?
2. What factors influence the development and evolution of teacher grading practices?
3. What changes (if any) might be implemented in current grading practices which would provide a better measure of assessing student achievement?

The chapter commences with a description of the setting and participants. The subsequent sections of the chapter trace the development of the major concepts and show how these contributed to the selective coding and the resulting grounded theory.

Introduction

According to the North Carolina Department of Public Instruction (NCDPI, n.d.), the Title 1 federal program exists to provide funding to LEAs with a high percentage of low-income families to ensure that students meet high academic standards. “Title I is designed to support state and local school reform efforts tied to challenging State academic standards in order to reinforce and amplify efforts to improve teaching and learning for students farthest from meeting State standards” (NCDPI, n.d., para. 4).

This study was conducted in three Title 1 middle schools in North Carolina. There are 11 middle schools in the district, four of which are classified as Title 1 middle

schools. These schools were targeted for the study in an effort to determine whether or not there is a need for improving grading practices as a form of Title 1 reform efforts. The academic year consists of two semesters, each of which is divided into two 9-week grading periods. At the middle school level, semester grades are derived from the average of two sets of 9-week term grades. The average of the two semester grades produces a student's final grade for the year.

Interim progress reports are sent home at the midway point of each 9-week grading period. The grades printed on these reports reflect student actual grades; however, at the end of each 9-week period when final grades are reported in each subject, the lowest grade teachers can report is 50%.

The district has established grading policies/guidelines to give teachers direction on expected grading practices. Some of the major grading guidelines provided for middle school teachers are summarized below.

1. Teachers are expected to enter grades each week through PowerTeacher.
PowerTeacher is an online grade book that enables students and teachers to access their grades.
2. Each teacher is required to have a syllabus approved by the administration and shared with students and parents.
3. The final grade for each 9-week grading period must be composed of a minimum of three tests, four independent assignments (quizzes, classwork, common assessments), and five guided work and homework, the frequency of which is determined by professional learning communities at each school.

Middle schools within the district follow a district created pacing guide which

outlines the content to be taught and tested within each 9-week grading period. Within this time, teachers must have the minimum required grades for each grading category (test, independent work, guided work, and homework).

Participants

According to Creswell (2012),

It is typical in qualitative research to study a few individuals or a few cases. This is because the overall ability of a researcher to provide an in-depth picture diminishes with the addition of each new individual or site ... because of the need to report details about each individual or site, the larger number of cases can become unwieldy and result in superficial perspectives. (p. 209)

In keeping with this general guideline from Creswell (2012), the study involved 14 participants which included 11 middle school teachers, one school administrator, one district leader, and one grading expert. Data collection took place over a 6-month period from May-November 2018.

The teachers who were interviewed represent various content areas and years of teaching. This is summarized in Appendix F. Pseudonyms have been used for all participants except the author, Ken O'Connor, who consented to having his name revealed in the study (see Appendix G for request to interview Ken O'Connor).

In grounded theory studies, data analysis occurs in three stages of coding: open coding, axial coding, and selective coding (Corbin & Strauss, 2015; Creswell & Poth, 2018). Transcripts of each interview were prepared and analyzed using memos and diagrams. During the initial phase of analysis, known as open coding, the researcher first read through the transcript several times to develop a general understanding of the

interview. After several initial readings of the transcript of the first interview, the open coding analysis began. This entailed breaking up the transcript into smaller pieces of raw data. Each segment of raw data was numbered and then analyzed using memos. Each corresponding memo was also numbered according to the raw data segment. The memos documented how the researcher was thinking while the data were being analyzed. This process was used continuously throughout the data collection phase. According to Corbin and Strauss (2015),

When doing analysis, researchers are interacting with data. They are examining it, making comparisons, asking questions, coming up with concepts to stand for meaning, and suggesting possible relationships between concepts. In other words, a dialogue is occurring in the mind of the researcher. Writing a concept in the margin doesn't preserve that dialogue or indicate how concepts might relate to each other. Memos and diagrams fill this role. (p. 107)

In accordance with guidelines from Corbin and Strauss (2015), each memo was captioned with a concept to represent the meaning the researcher ascribed to that segment of raw data.

During open coding, 51 different conceptual labels were identified from the first interview. These concepts were further developed during subsequent interviews, sorted and combined into more abstract categories and subcategories until data saturation was achieved. According to Corbin and Strauss (2015), saturation is achieved when “researchers are satisfied that they’ve acquired sufficient data to describe each category ... fully in terms of its properties and dimensions and that they’ve accounted for variation” (p. 240).

Categories

By the end of the study, three major categories were identified and developed from the data. These are presented below:

1. Grading practices (managing achievement and nonachievement factors)
2. Policies (the grading policies/guidelines that bind)
3. Grading Outcomes and opportunities for change

Research Questions

Data to address the research questions were derived from documents (syllabi, grading records, and district grading policies) and interviews with participants. All teacher participants were invited to tell their story of how they arrive at a final grade for each student. This entailed talking about the composition of a grade and the practices they use to assign grades to students.

When participants were asked to describe their grading practices, their descriptions all encompassed ways in which they manage achievement and nonachievement factors. They also described various reassessment strategies used when students fail to perform well on tests. Many participants described their struggle with grading student work products and grading the process by which students arrive at a final product. Some teachers have expressed that they place more emphasis on work that students produce independently, for instance, on tests and quizzes; however, they also value the process students use to accomplish those objectives. This is typically expressed by some teachers as work habits, showing work on a math test, citing textual evidence on a multiple-choice test, and showing effort.

The analysis from these interviews is presented below in relation to each research

question.

Research Question 1

What are the grading practices currently used by teachers? An examination of each teacher's grading records revealed that teachers organized grades based on assessment methods as opposed to standards taught. In general, the grades assigned by all teacher participants were based on predetermined categories of tasks dictated by the district's grading policies. All teachers are expected to assign grade(s) for homework which account for 10% of a student's overall grade, grades for guided work activities which account for 20% of a student's overall grade, grades for independent work which account for 30% of a student's overall grade, and grades for tests which account for 40% of their overall grade.

During each interview, participants shared the many unique practices they used when assigning grades to students. From these accounts, it was revealed that the approach taken by teachers in assigning these grades varies greatly from one teacher to the next and varies across schools, disciplines, and grade levels. All activities teachers engage in to assign grades were categorized as grading practices.

Grading practices. This category emerged from data provided by teacher participants, administrators, and documents (syllabi, grading records, and district grading policies).

Table 4 outlines the subcategories related to grading practices.

Table 4

Open Coding Analysis of Grading Practices

Category	Subcategory
GRADING PRACTICES (<i>managing achievement and nonachievement</i>)	The “E” grade Reassessment strategies The zero grade Homework

As participants described their grading strategies and habits, now categorized as grading practices, all their accounts seemed to encompass the various ways in which they manage nonachievement and achievement factors. Student effort was a commonly discussed topic among teachers. Many participants spoke about the various strategies they used to account for student effort in their grades. They also explained how they managed reassessment, despite its omission from the district’s grading policies for middle grades. In addition, homework and the use of zeros were frequently addressed during interviews. These recurring issues that were raised by participants led to the following subcategories related to grading practices:

1. The “E” grade (managing nonachievement factors)
2. Zeros
3. Reassessing students (managing achievement factors)
4. Homework

These subcategories are explained below in greater detail.

The “E” Grade

All participants seem to struggle with managing nonachievement and achievement factors when making grading decisions. The most prevalent nonachievement factor that emerged from the interviews was categorized as The E

Grade. This category represents all the data participants shared that were related to student effort and the part it plays in their grading. This seemed to be the most talked about issue and the one teachers seemed to wrestle with or defend in the most passionate way.

Parker and Tasha are math teachers at different schools. They both expressed how their grading practices change as the year progresses. They attribute this difference to the role effort plays at the beginning of the school year compared to other grading periods. In addition, they both shared their concern that if student grades were based solely on achievement, this would increase the failure rate in their classes and the likelihood of many students becoming demotivated. In an effort to manage this fear, these participants explained how they have used various strategies to mitigate this problem.

According to Parker, being strict in her grading methods means that she does not consider effort and other nonachievement factors when grading, but it also means that more students would be at risk for failing. To alleviate this problem, she adjusted her practices to include effort and work habits in order to help buttress student grades. In the statement below, Parker described her perception of the relationship between accurate grade reporting and student failure rate:

When I first started into teaching, I was very tough and I would probably say that my scores were pretty accurate compared to what I think they were going to do on the final assessment at the end of the year.... Now as I have grown the research says that retaining the child does not serve the purpose and that we're hurting their self-esteem ... so now I take into account other factors, not just the accuracy

but how they showed me work ... I get stricter in the fourth quarter and the third quarter with that accuracy.

According to Parker, she works hard to build good relationships with her students, especially in the beginning of the school year. One of the practices she uses to help establish this positive relationship with students is by awarding them credits for their effort and work habits like showing their thinking when solving math problems. Parker shared that she knows that this renders the grade less accurate but continues with this practice in order to avoid compromising her relationship with students. She explained that as the year progresses, her grading becomes more accurate and more aligned to the summative state assessments because effort is factored less frequently into her grades. She even went as far as to estimate that in the first half of the year, 60% of her grades are related to the process that students use to show their work or their thinking and 40% are related to a student's actual achievement or final work product. She further explained that towards the end of the year, there is a shift in her grading so that 60% of a student's grade reflects achievement or student work product and approximately 40% reflects the process (effort and other work habits) students use to achieve a certain task. This is expressed in her statements below:

In the beginning I don't know what they have learned before, you know, you hope that they have learned all the standards but sometimes that is not what happens so we need to build that relationship and build their confidence up ... so that's what I tried to do in my first quarter and second quarter before they are totally turned off to anything that I have to teach and ... instead of hurting their confidence I tried to build their self-esteem.... So sometimes I have to look at their work process at

the beginning.

During one interview, Tasha took a moment to log into her grade book in order to examine the grades she assigned to students. While doing this, she appeared to be comparing the grades she assigned to students at the beginning of the school year with their current level of performance. At the time of the interview, she had been working with these students for almost a full year and explained that she now has a better understanding of all her students and their capabilities. With this knowledge, she explained that there were a few students who were assigned grades in the first 9 weeks that, in retrospect, seem to be a poor reflection of their level of mastery or understanding. She was surprised to discover that one of her students who at the time of the interview was at risk for failing was previously assigned an 80% at the end of the first 9-week grading period. She explained that “he puts forth no effort ever, nine times out of ten he’s going to guess, as he even times tells me.” She continued the interview to explain that “his ability level, I think, could possibly be a good solid C ... but because he doesn’t put forth the effort, it’s inconsistent.” As she continued to scroll through her grade book and examine the grades assigned to other students, she appeared to have made a discovery:

Now that I’m thinking as I’m talking, I wonder if I subconsciously inflate the grades in the first nine weeks because I don’t want to totally deflate them ...

because if they start off extremely low then they’ll probably not work for me.

She continued by explaining how her grading changes from the first 9 weeks compared to other grading periods. “I’m probably more inclined to curve a test during the first nine weeks because I don’t want them to immediately feel defeated and then give up the rest

of the year.”

Several other participants shared the different conditions under which they assign grades to students in exchange for and to elicit greater effort from their students. Hanna, who teaches science, expressed the following during one interview: “If a student gives me some kind of effort – something – whether or not they are way off ... I will give you some credit for the effort because I don’t want them to feel completely defeated.” This statement from Hanna was typical of the overall sample in terms of the role that effort plays in assigning grades. Other teachers have expressed that when students behave favorably and comply with directions, they are more likely to assign extra credit to those students in order to encourage them to keep trying, even if their level of understanding and work do not correlate to the grade assigned. This is expressed in statements by John and Anders. During the interview, John explained how his grading differs from one student to the next because of his focus on student effort:

I grade basically on how well they’re at least attempting it and that’s at least a 50 for me. If you attempted it, and you put something down and you show me you’re using your time efficiently to do that, I will give you a 50 even if you get it all wrong.

Anders, who is a first-year teacher of English language arts, seemed to be very open, honest, and reflective during the interview. He explained that when he does not use a rubric to grade an assignment, he knows his grading is more subjective. He explained how a student’s effort may affect his/her grade. According to Anders, “I tend to be a little bit subjective ... especially if it’s not an assessment that is multiple choice and it’s not a clear-cut right or wrong answer.” He went on to further explain the ways in which

effort influences a student's grade:

When I see that she's working really hard or he didn't talk the whole class, but they just didn't do this correctly, I don't want them to fail or like feel unsuccessful because there's kids in that class that didn't do anything but scribbled something down and just happened to be right or maybe they got it from someone else. I just don't feel ok with them getting a higher grade than the kid that actually tried ... I let behavior influence it a little bit so the kid who worked really hard, I always give them a better score.

Anders further explained his reasons for considering effort in his grades when he said the following:

On an assignment where it's just, read this and tell me your thoughts, that's hard to get a low grade as long as you read and actually gave me your thoughts (pause), so in that way they're rewarded for just attempting the work ...but I feel if I don't grade things like that they just don't do it ... they'll say Mr. Anders, "Is this being graded?" If I say no, they don't do it.

Some of the phrases or words frequently used by participants and which are associated with this category are

- He really tried hard.
- She works really hard.
- They did everything you ask them to do.
- She turned in all the work.
- They always participate.
- She does everything you ask them to do.

The Zero Grade

All participants described how they handled missing work or the conditions under which they assign a grade of zero. Hanna was very passionate when she spoke about the zero grade. “A student will get a zero from me ... if you just don’t do the assignment at all, you make no effort whatsoever to do the assignment, you are getting a zero. Period.” She continued the interview by clarifying that she does not want students to give up, so once they show some level of effort, they will receive credit for that work. She expressed deep concern for her students and the effect a zero can have on their level of motivation. This is reflected in her statement below:

I want them to keep trying and a lot of my students I have to be careful with that because if they feel like they just can’t do it ... they get a zero every time they try to do it, they’ll just stop. They won’t make any effort at all.

Parker explained that once a student takes steps to make amends and shows that they care, she works with students to provide them with an opportunity to redo an assignment; however, she explained,

If a student is just being rebellious and is not completing the work ... I say well to me we have gone through this and you have been taught several of this, you know several items on here so the least you can do is underline and circle meaningful statements and word problems and if they’re just being defiant then it’s a zero. Now does that zero stay? No ... maybe they had a bad day they can always come and talk to me and they can always redo that assignment.

Tasha and Gad both mentioned that a student will get a zero when they turn in work without writing their name on the paper. They explained that when this happens, it

creates a problem when they are trying to figure out who the work belongs to out of a set of 70 or 80 pieces of work. This is expressed in the following statement from Gad:

In most cases a student will earn a zero grade if they don't put their name on a paper, if they turn in a blank paper and in most cases, they get a zero when they just don't do the work.

Gad explained that he makes every effort to call parents and provide extra copies of the assignment; but there are those students who do not respond to any strategy a teacher uses to get them to do the work, so a zero grade remains on their report card for the grading period.

According to Sally, "Some of my high flyers don't have grades that represent how well they could be doing. Some of them simply lack the motivation ... I'm not going to chase 12-year-olds down to get their work." She continued to explain that under those circumstances, she assigns the lowest possible grade allowed at her school when students do not make an effort to turn the work in. She explained that the lowest grade that any teacher can assign is a 50% even when work is missing. She was the only participant who reported that missing work was assigned a minimum grade of 50%; all other participants reported using zeros for work that is not turned in. The policy at her school prohibits teachers from assigning a grade of zero.

A common concern expressed by several participants was how to keep students motivated to do the work so that they can avoid getting zeros. A frustrated Hanna explained during the interview that the biggest challenge she encounters in any grading period is "just getting them to do their work." Six other participants including Gad also expressed that this was one of their main struggles with grading. "I'm not going to give

them a grade when they give one-word answers. I expect that in middle school you should care more to do your best work.” Gad continued the interview by explaining that “The student might know the content, but my job is to see it on paper. I don’t give oral quizzes. I am preparing students for life. So, when the work is missing, you get a zero.”

Homework

Homework accounts for 10% of a student’s grade as determined by the county’s grading policies. Participants have expressed the various ways in which they manage this. Overall, the teachers who were interviewed all demonstrated some level of dislike as it pertains to the assignment of homework grades.

According to Tasha, during the first 9 weeks, she assigns more homework to students compared to other grading periods:

I assign more homework to students compared to other grading periods.

I enter homework grades on a weekly basis. If you completed it, you got a hundred, if you did most of it you got a 90, if you did maybe ½ of it you got an 85% and if you didn’t do any of it I didn’t want to completely bomb you out because homework is one of those things that has to be reinforced at home, so I’ll give you like a 70. I don’t put it in as a 0.

She explained that since she has no control over what happens at home, she did not want the zero to affect a student’s grade even though homework is only 10% of the grade.

Tasha, who previously taught at an elementary school, explained that she hesitates to assign a zero for homework because in elementary school they were not allowed to assign a grade for homework, and this was a change for her in moving to middle school. She still believes that home influences should not affect a child’s grades, so she does not

assign zeros when homework is not completed. Other teachers have shared different strategies for managing homework.

When asked to describe how homework is graded, John's response was swift, "Oh I don't give it here." He explained that he simply assigns short classwork assignments to students and explains to them that it accounts for 10% of their grade. When asked why he does not assign homework, he had this to say:

I just think because of the socioeconomic status of our kids, the environment we're in, I believe ... that all learning happens in the classroom. I've given homework at this school and I've gotten 5 or 6 back and of the ones that were turned in maybe only 2 items were completely right.

He also continued by explaining that there are so many students who do not have the technology at home and that it would be needed for them to complete most of his assignments. Four other participants reported that they handle homework in a similar way.

Hanna spoke passionately about homework as she explained why she does not value this grade and why she hardly assigns any work in this category. During the interview, she explained the following:

I tend not to give a lot of homework, deliberately, because of the whole fidelity think. They're just cheating off of each other. I have a hand-full of kids who will do it all by themselves. They're always my high flyers. They always get the highest grades in the class; they're always the kids that everybody is going to get the answers from, and you can walk in their homeroom and watch them copying off each other. It's pathetic.

She continued the interview with, “the only reason I give homework is because I’m forced to by the county!”

A few years ago, Hanna was invited to serve on a committee which would discuss and make changes to the district’s grading policies. The committee which met during the 2010-2011 school year consisted of teacher representatives from all schools in the district. The purpose of the committee was to establish grading guidelines for the district. This information was corroborated by two district administrators. Hannah spoke with pride when she explained her role in helping to establish the current grading practices and, with equal pride and passion, when she explained how she had to advocate for her students at this meeting. She explained that had it not been for teachers like her, the policy on homework would have been more severe. She explained that the topic of homework raised a lot of questions and was the source of several heated discussions. According to Hanna, the topic of homework was

a huge sticking point that was a pretty in-depth conversation with the entire group because you have one representative from every middle school in the county and they had some pretty strong opinions because you have those middle schools that have helicopter parents. Those parents expect their children to bring homework home; they want to know why their children are not coming home with homework; they want you to give their children more homework because they feel like if they’re piled up with homework, they must be learning a lot. Right?

She then went on to explain that the debate on how to factor homework into the grades was one that was argued between teachers from Title 1 schools and those from other non-Title 1 schools:

It literally boiled down to the Title 1 middle schools saying guys ... I understand y'all's predicament, but you have to understand ours.... How do I look at a kid who probably had to sleep in a car last night, who's not quite sure where he is sleeping tonight, he's not sure if his clothes are going to get washed ... I literally had a child one year (and I told them this), who came home and there was a lock on their door, and they couldn't get into their house. All of this child's belongings were in this house and she could not have them. All of her clothes, all of her toys, all of her mementos, everything and I said, so she comes home and I'm going to ask her to sit down and do a writing assignment. Really? They're not sure where they're going tonight let alone what's going to happen now. I said this happens to our kids all too frequently.

She continued the interview by explaining how she kept up her fight to ensure that homework grades would not have a huge impact on a student's overall grade.

When faced with those situations, she tells her kids that "I'm just glad you're here today, here's a pencil, here's a notebook ... what other school supplies do you need?" She explains that those things become a priority in Title 1 schools. She explained that teachers from other schools with a more affluent student population did not understand the differences in their worlds. She explained that the policy now requires that teachers report grades for at least five guided assignments, four individual assignments, and three tests. "You cannot have below those numbers; that's a middle school requirement and they wanted to set a number for homework." She continued,

For homework, I could literally give one homework assignment and I would be fine. We had to fight for that and thankfully the particular superintendent that

was in charge of that understood ... and knew what we were talking about because he was over there agreeing with us and nodding his head the whole time and smiling, and these other teachers who have never worked in a Title 1 school didn't understand.

Most teachers who were interviewed seemed to share a similar sentiment as it pertains to homework. Most teachers expressed their frustration with the requirement that homework must be included in a student's overall grade. Their main frustration seemed to stem from student failure to complete homework.

According to Bailey, she would eliminate homework if she could because she seldom assigns homework to her students. "It's very few and far between assignments that I actually grade for homework." She further explained why she hardly assigns any homework:

I haven't found any great research that says that homework does benefit test scores at the end of the year and if they take something home and are doing it wrong, when they come into my classroom, I have to fix everything they've done wrong for the last hour at home, so it just doesn't make sense to me.

She continued the interview with the following:

I don't like taking completion grades just because if they're gonna do something I feel as though I should grade it and grade it with fidelity. I'm not just gonna give you a hundred if you put something down on a piece of paper because that doesn't show me that you understand something.

Similarly, Anders shared lessons he learned from a graduate school class he is now taking. "They say homework is more detrimental because it takes up your time trying to

fix it and then the kids don't do it half the time."

Larry also shared that he is really struggling with some of the homework requirements in language arts. He explained that for homework, students must read for 30 minutes each night:

Do you really want them to read 25-30 minutes at home every night? How do we get them to do that because we have some parents who work the night shifts and can't initial what their students read, so the accountability piece of that is very hard to figure in, along with, okay well this student took home a lower level text but he still reads it even though that's not on his level, do I count that or do I not?

According to Linda, if she could change one thing, it would be to eliminate homework requirements:

Homework could possibly be eliminated because ... I'm not sure if they're doing it or its someone else at home doing it. Sometimes someone at home may be doing it with them ... and it doesn't really reflect what they can really do. When they're in class we can see what they can really do. Many times, in one of my classes now, we're doing homework in class anyway because many of them don't do it and many of them don't have access to technology that is sometimes needed.

She continued to explain that many of the zeros students have can be attributed to their failure to complete homework. She thinks students would be more encouraged if there were no homework assignments they were obligated to do. She believes it would be less burdensome for the students. "I think it's less stressful for them to think I have to do this homework and they may never get this done."

When Sally was asked about homework, she quickly replied, "I do not grade

homework.” She then continued by explaining,

This is a transient society within this school population. A lot of these kids honestly cannot go home and do homework. For many reasons: babysitting, taking care of siblings, no access to Internet, or simply not just having the supplies at home they need. So, I just do not do homework.

When asked to explain how she works around this, given that homework is required, she explained how she makes allowances for her students in class:

The way I do it is that I may start homework at the end of class and, when they come in the next day, I may give them 8 to 10 minutes to complete it. That is your homework time.

Reassessments/Retakes

Participants all had different strategies for managing retakes associated with test grades. A common strategy that was described by several participants was the use of test corrections. There were variations in the way participants handled grade calculations after test corrections/retakes are done.

The variations included averaging the two scores; awarding students with a grade that is 50% higher; and, in a few cases, giving students the higher of the two grades. Hannah explained that she wants to ensure that the retake is done “with fidelity”, so a student has to make an appointment with her so she is able to sit with the student to learn more about their misconceptions. She explained that she values this method because it gives her feedback on her teaching and gives her greater insight on what other students might be struggling with. She places emphasis on the corrections being done in her presence, so she can verify that they are doing it correctly and without help from parents

or other students. At the end of a session, the student is rewarded with half credit more.

She explained,

They cannot do test corrections on their own. They can earn up to half credit back ... let's say they made a 50 ... after test corrections they earn an extra 25 points and so their grade is now 75 which is a whole lot better than a 50.

Parker, however, explained that she does not require students to do the test corrections in her presence:

They can do it and take it to me.... And I know I'm vulnerable for like a big brother or a parent or a neighbor to do it.... But if they're going to take that incentive to do it, I'm going to give them credit ... and once again that's part of building that relationship of trust.

Parker further explained that she does not want to discourage them or provide any additional hoops they have to go through to get retakes done. She wants to motivate them to make the corrections, so she allows them to do it on their own and then turn in the corrected assignment to her. She described an interesting approach to grading retakes:

All my students can redo an assignment once and what I do then is average the two scores ... if they get a 90 on the assignment and they're going to take the time to better themselves ... and take that step to understand that problem ... or the process, then I will average that grade and the child will get, if they get a 90 again that it's a 90; if they get a 96 it will be a 93; whatever that child gets I average the scores.

Sally explained that she grades exit tickets as quizzes. She explained that the kids

would have a three question exit ticket at the end of class for each concept ... normally every 4 days. On that exit ticket they know, the most I can miss is one. If I miss more than one, then they know they would not get the opportunity to make this up. That gears them to really pay attention, because they want good grades from the exit ticket.

She further explained that she makes allowances on tests since they account for 40% of a student's final grade. She also allows students to do test corrections: "When kids don't show mastery, they have an opportunity to take that same test, go back and make corrections. But they only get half the points back."

Gad also had an interesting approach to retakes. He explained that he does not publicize any retake policies on his syllabus because he wants students to try really hard the first time. Instead, he will go to the individual student to encourage them to try again. Parker seems to value student effort the first time around too. She explained why she averages student retest scores: "It benefits the child that they can up their scores after a retake, but it doesn't put them over the student who came and got a hundred the first time." Similarly, Baily puts a limit on the highest grade students can make after they make test corrections. She explained her reason for doing this: "It's not fair to the students who got it the first time. I don't want to put them over the kid who did it on the first try, so the highest they can make is a 60." Larry and Anders also utilize test corrections, but they give students extra credits after corrections are made.

According to James, for students who do fail the test, "I do give them the option to retake it and they'll get an average of the two scores. I've had maybe three or four kids in two years to do it." He then continued by explaining how he manages retakes

differently since most students do not take him up on the offer. “For the most part, I curve my tests.” When asked to explain what he meant, he explained how he uses the state’s grading scale on the final summative assessment for social studies as a guide. He explained, “you can get like a 30 and still get a D, just because of the way that scale works.” He uses this scale to convert or curve student test scores.

Table 5 shows the categories, labels assigned to these categories, and the properties and dimensions associated with grading practices.

Table 5

Properties and Dimensions of Grading Practices

Subcategories	Labels	Properties	Dimension
The “E” grade	Working hard, effort, trying,	ability engagement student effort	Low-high
The zero grade	Missing work, turning in work		0-50%
Reassessments	Test corrections, retake	Extra credit, grades averaged,	Used – not used
Homework	Homework	Work done outside the classroom	Assigned – not assigned

Research Question 2

What factors influence the development and evolution of teacher grading practices? When participants were invited to talk about influences on their grading practices, most participants explained that their grading activities were largely determined by the existing school or district policies. Some of the other influences included the following:

- College education courses
- Colleagues/mentors
- Personal experiences as a student
- Student reactions to my grading methods
- School grading policies
- District grading policies.

The most common influence teachers described was related to district policies.

This information provided by teachers during interviews prompted the researcher to conduct theoretical sampling in order to obtain more data to better understand the policies that influence teacher grading activities.

Policies (The Policies and Practices that Bind)

This category emerged from data obtained through interviews and theoretical sampling of documents. District grading policy was a topic that was addressed by several participants, both explicitly and implicitly, as a major influence on teacher grading and was later selected as the central phenomenon in accordance with the guidelines from Creswell and Poth (2018):

This is an aspect of axial coding and the formation of the visual theory, model, or paradigm. In open coding, the researcher chooses a central category around which to develop the theory by examining his or her open coding categories and selecting one that holds the most conceptual interest, is most frequently discussed by participants in the study, and is most “saturated” with information. The researcher then places it at the center of his or her grounded theory model and labels it central phenomenon. (p. 316)

Table 6

Open Coding Analysis of Policies

Category	Subcategory
POLICIES	Expectations and Intent
	Grade composition
	Documentation/communication

An analysis of the grading policy document together with data provided mainly by one district leader yielded the subcategories outlined in Table 6. These are explained in greater detail below.

Expectations and Intent

According to the policy, its purpose is to “provide direction to all teachers in expected grading practices.” A summary of the grading expectations as listed on the website is as follows:

Elementary, middle and senior high grades reflect student performance on the various subjects (such as English, math, etc.) at each grade level in the North Carolina Standard Course of Study. Student grades are based on tests, quizzes, projects, assignments, homework, performances, etc. with different percentages. Grades for assignments are based on the importance and time spent on the activity.

In addition, four specific expectations were outlined on the first page of the grading policy. These are shown below as they appear in the document:

- Grading practices shall be reviewed annually by the principal and School Improvement Team.
- Academic grades shall be a content-based measure of what students are able

to demonstrate.

- Grading shall provide appropriate and accurate feedback to the students and parents.
- Grades shall reflect performance that is consistent with demonstrated achievement on the goals of the North Carolina Standard Course of Study.

During an interview with one district leader, Dr. Bentley, clarification on each of these expectations was sought and provided.

Grading practices shall be reviewed annually. According to Dr. Bentley, this expectation is monitored on the district level by the three executive directors who oversee elementary, middle grades, and high school. She explained that “they meet monthly with the school principal to review the school improvement plan which includes the grading policy and practices.” She went on further to explain that “the purpose of reviewing the policy and practices is to make sure that we don’t have grade inflation ... so that what we’re giving students is a reflection of their work in the classroom.” Dr. Bentley explained that the intent is for the executive directors to “ask probing questions of the principals to get them thinking.... So we are trying to provide some fidelity checks throughout the district to make sure that it is cohesive from grades 3 through 12.”

On the contrary, one administrator who was interviewed confirmed that the school improvement team does not examine grading practices each year and expressed regret that this expectation has not been carried out. She explained that the school simply follows the policy laid out at the district level. The teachers who were interviewed also reported that their school improvement team did not meet to discuss grading policies each year.

Grades shall be content-based measures. Dr. Bentley continued the interview by explaining that the goal of the “county and most districts is to move to mastery of concepts ... assessing the child to see if they’ve mastered the content.” The researcher sought clarification on what was meant by content-based measures or mastery of concepts and was told that this is synonymous with mastery of standards. The policy makes frequent reference of content-based measures which can be interpreted as being equivalent to standards-based mastery.

Grades shall provide accurate feedback. When probed about the level of feedback the district hopes to communicate, Dr. Bentley spoke about the existence of a parent portal and other systems of communication: “We have the portal now and so parents can immediately see grades.” In addition, she explained that the chief accountability officer and their teams have also taken this a step further by doing the following: “They go in and put the percentages of grades, what the grades are worth” to help provide “accurate feedback to give an appropriate reflection of how the student is doing in class.” She continued by providing the researcher with more background information that would help explain why these policies were created: “So, for example, in the past, homework may have counted for a higher percentage than it should have, so you’re giving a parent a bad or inaccurate reflection of if the child is mastering the content in the classroom.”

Dr. Bentley further explained how past inconsistencies among teachers provided the impetus for items listed in the grading policy. In reference to homework assigned by teachers, she explained, “sometimes for homework, if you try it, you’re gonna get credit for it ... so we’re really trying to do a better job with our percentages, so they are a better

reflection of ... really rigorous instruction.” Furthermore, Dr. Bentley explained that with this, teachers will be better able to provide feedback to parents as to whether or not “your child is either being successful or they’re not and this is how we can help them be successful in that area.”

Grades shall reflect performance on the NCSCOS. The final expectation listed on the first page of the policy addressed the relationship between grades and the Standard Course of Study. When asked to explain this, Dr. Bentley made reference to grade inflation and the work the district leaders do to minimize or eliminate this:

What we were finding when we were looking at things to predict how a school was gonna do, how a teacher was gonna do, a child could have an A in biology in high school and when they take the EOC they’d have a one. So that said to us that we did not align our pacing and instruction, we didn’t hold the teacher to the standard that the major portion of their assessment on a student ... should be on the unit or chapter test so that we have a better idea, is this child gonna be successful on the EOC.

She continued by explaining,

It’s educational malpractice, we felt, on us if we allowed a child all year to think they were doing what they needed to and get to the state assessment (which should have been what we tied our standards to, or assessments to) and they weren’t successful.

In summary, the overall expectations of the grading policy are that principals and school improvement teams review them annually, they should be aligned to state standards, and they should provide accurate feedback to students and parents.

Grade Composition

When teachers were asked to tell their story of how they arrive at a final grade, most of their responses centered on the grade composition outlined in the district policy. The purpose of this section of the district policy, as stated on page 2 of the document, is to ensure the following:

Grades shall reflect student performance in various subject areas aligned with the North Carolina Standard Course of Study. The make-up of the student grade shall be based on various assignments with differing percentages. The assignments shall carry different weights based on the importance of the activity, time spent on the activity and the degree of involvement.

Data gathered from the district website also reveal that prior to the 2010-2011 school year, there were no district grading policies in existence. The information found on the district website revealed that the district,

like many other districts in the state, had no countywide grading guidelines or practices that outlined how much a quiz, homework or a major test (six weeks test) would count. Nor did the district have guidelines for the minimum number of tests that must be given in a course. Therefore, it was left to each school or the classroom teacher to decide whether to give one or six tests and the value of a test as compared to homework, class work or a special project assignment.

In addition to the data obtained from the district website and the grading policy, additional context was provided through the interview with Dr. Bentley who explained that the overall purpose of specifying the composition of a grade is to “support the teacher so that they’re not on an island, but also to support the students so that they’re getting the most rigorous instruction wherever they are in the district.” She explained

this in greater detail:

There is a progression as you get from elementary to middle and to high with the expectations, but to protect the integrity of the diploma with so many high schools and middle schools and 32,000 students, we need to make sure that if we're giving a test in high school, that it is consistent with fidelity across the district so that if you happened to end up at this high school and this teacher doesn't have the same expectations then you're apt to do better on a test, but if you get a state standardized assessment, is it going to be a true reflection?... Again, going back to the grade inflation.

She continued the interview by expounding on the supports the district offers as a means of guaranteeing the integrity of the grade:

We have a lot of our NC finals and a lot of things that are standardized across the district, even some chapter assessments, unit guides; so, when you have all these things in place, you're more apt to get fidelity. So those are the things we were measuring students on to be consistent across the district.

The district grading policy requires that teachers assign grades for homework, guided work, independent work, and tests. The grades assigned by all teacher participants are based on predetermined categories of tasks dictated by the district's grading policies. Teachers must assign grade(s) for homework which account for 10% of a student's overall grade, grades for guided work activities which account for 20% of a student's overall grade, grades for independent work which account for 30% of a student's overall grade, and grades for tests which account for 40% of their overall grade. Additionally, the policy states,

A minimum of three (3) tests shall be given during each grading period for all middle school courses. In addition to test grades, a minimum of five (5) guided assignments and four (4) independent assignments shall be given in each grading period for all middle school courses.

Documentation/Communication

Various forms of documentation and communication were emphasized throughout the grading policy and during conversations with participants. Table 7 represents the grading documentation required by the district at each level. Table 7 shows that at all school levels, the district requires communication between teachers and students/parents about their grading plan. In elementary schools, the grading plan must be created by the grade level team and for middle and high schools, each teacher must express this in the form of a syllabus.

Table 7

Required Grading Documentation

Elementary	Middle	High Schools
Each grade level within a school shall be responsible for providing a written document to parents explaining their grading guidelines.	All teachers are required to have a syllabus for each course taught.	All teachers shall be required to give the Principal or his designee and students a syllabus that contains information on how a grade is achieved. Principal shall review annually.
Each grade level's document must be approved by the principal.	The syllabus must be approved annually by the administration.	
Upon approval, the grade level document shall be sent home and signed by parents within the first two weeks of school.	The syllabus must be communicated annually to students and parents via website and/or handout.	
Teachers shall place the signed document in the Student Accountability Folder each year and purge at the end of the year.	The syllabus must include the breakdown of grading plans to include percentages or points.	
This document shall be given to all students at their time of enrollment.	The syllabus must include a list of classroom procedures.	

Additionally, the grading policy requires the following:

- Teachers shall supply a copy of their grading procedures to the principal, students, and parents.
- Teachers shall enter grades in PowerTeacher weekly.

School Policies

There were some variations in the way teachers grade based on policies that are in place at the school level. For instance, Gad, Parker, and Linda have shared that they are

required to offer students retake opportunities, but it has to take place before school or after school. They explained that their administrator is opposed to using instructional time to provide reteaching/retesting opportunities to students. They expressed that they wish they could have more flexibility with this because students do not seem to take the offer when it is after school or earlier in the day. Some have said this presents a challenge for students who do not have transportation to get home.

Other teachers simply accommodate the student when they make requests for retakes, or they simply allow the student to make test corrections.

Sally explained that at her school, she is not allowed to assign zeros anywhere. The lowest score a student can get on any individual assignment is 50%. Teachers at other schools assign zeros on individual assignments; but when reporting a student's overall grade at the end of a quarter, the grade must be at least 50%.

The main sections of the policy that give direction to teachers on the expected grading practices were summarized under the categories intent, composition, and communication. Together, these help to frame the approaches teachers use when summarizing a student's level of achievement with grades.

The intent, as listed in the grading policy, is to ensure clear communication so students know how they are progressing. The grade composition, as explained by Dr. Bentley, was included in the policy as a safeguard to prevent grade inflation. Additionally, teachers are required to keep parents and students informed about the makeup of a grade and how students are progressing.

Given that the overall intent of the district policy is to ensure grades provide accurate feedback to students that is consistent with the North Carolina Standard Course

of Study, what opportunities exist for changing and improving the existing policy?

Research Question 3

What changes (if any) might be implemented in current grading practices which would provide a better measure of assessing student achievement?

Data from interviews with teachers and district leaders/administrators suggest that current grading practices are, to a large extent, dictated by the existing district policies. Teachers must operate within the confines of this policy. The third research question was therefore resolved, in part, by analyzing and evaluating the items listed in the district grading policy. Accordingly, data to evaluate the policy were gathered through interviews and consultation with an expert in the field of grading.

The final participant to be interviewed was Ken O'Connor, an expert on grading and reporting as well as the author of several books and articles related to grading and effective communication of student achievement. Prior to conducting the interview, a copy of the district's grading policy was provided for his review. Since the study was focused on middle schools, his review and responses to questions posed were limited to the policies pertaining to middle grades. As a leading expert on the subject of grading, O'Connor provided data and insight to help resolve the third research question.

A Long Talk with Grading Expert and Author Ken O'Connor

The interview began with the following question being posed to the author: "What was your overall impression of the district's grading policy?" His response to this opening question was both quick and surprising: "Without exception, this is one of the worst grading policies I have ever seen!" We then spent the next hour discussing, in

detail, why he had drawn such a blunt and scathing conclusion. The interview protocol used for the interview with Ken O'Connor can be found in Appendix F.

The following represents data from this interview that address the third research question: "What changes (if any) might be implemented in current grading practices which would provide a better measure of assessing student achievement?" The complete grading policy document used by the district is listed in Appendix H.

An Evaluation of One District's Grading Policy (The Policy, Problems, and Opportunities for Change)

Written documentation. The first item in the grading policy addresses the written documentation teachers are required to have and share with parents. The main document that the district requires is a syllabus. Copies of these were shared with the researcher during the interviews with teachers as a means of triangulating data. Table 8 shows the five middle school grading documentation expectations that must be addressed with a syllabus and O'Connor's evaluation of each.

Table 8

Evaluation of the Grading Documentation Expectations by Ken O'Connor

Grading Documentation	Comments from O'Connor
All teachers are required to have a syllabus for each course taught.	Good.
The syllabus must be approved annually by the administration.	Good.
The syllabus must be communicated annually to students and parents via website and/or handout.	Good.
The syllabus must include the breakdown of grading plan to include percentages or points.	Not acceptable. This item needs to be changed to reflect levels of proficiency.
The syllabus must include a list of classroom procedures.	Good.

Table 8 shows that all, except one, of the documents required by the district were found to be acceptable. Since the district policy requires that grades be aligned to standards, O'Connor asserted that the grading plan should not include percentages but, instead, levels of proficiency. "If we're working from standards it shouldn't be percentages or points, it should be levels of proficiency, somewhere from between 2 to 7 levels." This was explained in greater detail throughout the interview.

Grade composition (percent makeup). The second page of the grading policy addresses the makeup of a student's grade. The first item for middle grades 6-8 under this section is:

The 9 weeks grade shall be determined by the following:

- 40 percent – tests (major – projects, papers and performances)
- 30 percent independent work – (quizzes, classwork, common assessments,

etc.)

- 20 percent guided work – (group activities, teacher led activities, etc.) and
- 10 percent homework

O'Connor's reaction to the items listed in this part of the document was a passionate one.

In reference to the grade composition, he strongly expressed the following: "It is completely the wrong basis for determining grades ... grades shouldn't be determined by assessment methods or activities" but, instead, "grades should be determined by standards." In addition, O'Connor stated that homework should not be counted as a grade as the purpose of assigning homework is to provide students with practice.

Grade composition (number of assessments). *"The single worst thing."* The second item for middle grades 6-8 under this section, grade composition, pertains to the number of assessments:

A minimum of three (3) tests shall be given during each grading period for all middle school courses. In addition to test grades, a minimum of five (5) guided assignments and four (4) independent assignments shall be given in each grading period for all middle school courses.

This, according to O'Connor, is "probably the single worst thing" in the grading policy. O'Connor went on to further explain: "This is the perfect example of the notion that school is about the accumulation of points. The more points you get, the better your grade, whether it has anything to do with learning or not." He explained that this should be changed because the focus needs to be on learning and this policy hinders students from going through a proper learning cycle. This was expressed clearly when he said,

You cannot teach for learning if you have that number of assessments ... you

have to have a teaching learning process where you do some initial instruction, kids learn, they practice, they get feedback on what they're doing ... and preferably several times, then we have an assessment that counts as part of a grade.

Band and basketball. According to O'Connor, when so many assessments are required, students do not have sufficient time to practice. He uses the following analogy to highlight the importance of practice as part of the learning process:

Students understand that in band and basketball practice counts. They need to have that same understanding in the classroom that practice counts in the sense of building the skill, building the understanding, building the knowledge before you have the opportunity to demonstrate. In band and basketball, you practice, you perform in a concert, you play the game. You have lots of opportunities to practice before the performance takes place. With this setup there is no opportunity to practice before the performance, before the demonstration of achievement.

O'Connor used this analogy to demonstrate that the grading policy requirement of a minimum of 12 assessments per 9-week period hinders students from experiencing sufficient practice before a performance grade is given. He continued, "The idea that in a 9-week period, you would have 12 assessments that count as part of the grade is just completely and entirely insane!"

Grade composition (other requirements). Other guidelines listed under this section, were regarded by O'Connor as being acceptable, with the exception of one item which required teachers to enter grades in PowerTeacher on a weekly basis. The

following items were given a favorable approval rating by O'Connor:

1. Academic grades shall not be reduced for misconduct or for punishment.
2. Teachers shall supply a copy of their grading procedures to the principal, students, and parents.
3. Teachers shall work with Professional Learning Communities to determine interventions and monitor progress to ensure mastery of content knowledge.
4. School level professional learning communities shall help determine an appropriate number of additional assignments to be given for each discipline.
5. The teacher shall adhere to the IEP, the 504, and any Tier Plan.

With respect to the requirement that teachers enter grades in PowerTeacher weekly, O'Connor clearly explained why increasing the number of assessments does not render the grade more valid. During the interview, the researcher also shared with O'Connor that a few teacher participants expressed that this requirement was closely monitored by the principal. According to reports from Gad, Parker, and Linda, teachers at their school were told that this expectation will help to ensure that the grades will be more accurate. In response to this, O'Connor stated, "this is absolutely false." He continued,

Assessment is a sampling procedure. It's about having enough evidence to make valid and reliable judgments. Not too much or too little. The measurement experts say that to make a judgment about anything we need a minimum of three pieces of evidence.... So ideally, we would maybe need 4 or 5 but we certainly don't need 14 or 24 or 34 pieces of evidence.

Opportunities for Change

Does this count for a grade? During the interview, the researcher shared a concern that many participants seemed to be grappling with. Participants expressed that if they do not grade everything, students do not take it seriously. In response to this, O'Connor stated that he is not surprised by student responses, because the grading policy encourages this type of response. He explained that the existing policy: "Trains students into the 'does this count syndrome' because there is so much that counts, so students don't have time to do anything that doesn't count.... It's the procedures that are establishing the 'does this count syndrome' in students."

O'Connor explained that the items listed under grade composition presented the greatest opportunity for change that would help make grading practices more effective and reflective of student achievement/learning.

Homework assignments. Page 3 of the grading policy lists five expectations for homework assignments pertaining to middle grades which are shown below:

- Homework assignments shall be designed to provide practice, preparation, or extension to curriculum taught in the North Carolina Standard Course of Study.
- Teachers shall work with Professional Learning Communities to determine amount and frequency.
- Teachers shall assign a reasonable amount of homework and it shall count no more than 10 percent of the final grade.
- Students shall receive feedback on homework.
- Teachers shall avoid using homework as punishment.

Table 9 summarizes the evaluation of the homework policy.

Table 9

Ken O'Connor's Summary Evaluation of the District's Homework Policy

Homework Assignment Policy	Comments from O'Connor
Homework assignments shall be designed to provide practice, preparation, or extension to curriculum taught in the North Carolina Standard Course of Study.	Good
Teachers shall work with Professional Learning Communities to determine amount and frequency.	Good
Teachers shall assign a reasonable amount of homework and it shall count no more than 10 percent of the final grade.	This policy would be ok if it specifies "extension homework" and if the 10% requirement is
Students shall receive feedback on homework.	Good
Teachers shall avoid using homework as punishment.	Good

In response to these expectations, O'Connor had this to say:

The first one under homework, is good ... but ... the third requirement which states that homework shall be no more than 10% of the grade, is not acknowledging that practice has no place in grades, preparation has no place in grades; extension may have a place in grades.

He then further explained that if the policy had specified "*extension* homework shall count for no more than 10% of the grade, then that would be ok, because extension generally is summative assessment." In addition, he reinforced what many teachers also expressed about homework.

There are some caveats to that ... we have to monitor it carefully to ensure that

it's the kids' own work and that it wasn't done by Dad or Mom ... and we have to monitor it carefully to make sure that all the kids have the resources to do it outside of the classroom. If those conditions apply and it is extension, then it's legitimately part of the grade.

He continued the interview by explaining that even if all those conditions were to be met, the 10% requirement would still pose a problem:

I wouldn't put a fixed limit on it because if you had 1 in every 9-week period, then 10% might be too much. If you had the kids doing it 2 or 3 times, then 10% wouldn't be enough. So again, it's the fixed percents that are problematic.

Interim/report card grades. Page 4 of the policy addressed report card grades.

There are four expectations listed:

- Teachers shall not give a grade below 50 on the report card for all marking periods.
- The teacher shall utilize the appropriate PowerTeacher comment on the report card.
- Interim reports shall reflect actual grades.
- The teacher shall maintain frequent communication with student and parent.

O'Connor commented on the first and third expectation. In response the expectation requiring that all grades be at least 50%, O'Connor agreed with the stated purpose, that all students should be given the opportunity to earn a passing grade, and expressed the following: "I agree that there should never be anything below a 50 if we are using percentages. But we shouldn't be using percentages."

With respect to interim grades, it was explained to O'Connor that these are given

at a midway point within each 9-week period and is sometimes referred to by teachers as a progress report. In response to this, O'Connor explained that

3 or 4 weeks is probably too soon for a progress report but if it's a progress report it should not have grades ... if it has grades on it, then it's not a progress report, it's an achievement report.... A progress report would have a completely different system of labeling.... A progress report after 3, 4 or 5 weeks, all it should say is "*on track*" or "*not on track*."

Core/encore classes. This section listed five guidelines which are listed below:

- All Middle school courses shall carry the same graded weight regardless of the subject or academic level.
- All middle school courses shall establish procedures for grading and assessing their students. This includes but is not limited to Health and PE, Band, Chorus, Art, Foreign Language and CTE courses.
- The grades shall reflect what a student knows and is able to do.
- All classes shall use a variety of assessments to measure learning.
- Tests shall include but are not limited to performances, physical activity assessments, content assessments and performance-based projects.

O'Connor found no problems with this section and regarded this portion as being more related to philosophy than grading practices. He embraced the guideline which encourages the use of a variety of assessments, stating, "Anything that suggests that variety is necessary is a good thing."

Reteach/retest. There were no guidelines for middle grades under this category. O'Connor's evaluation is shown in Table 10. O'Connor took a moment to comment on

the overall expectations listed on page 1 of the policy:

- Grading practices shall be reviewed annually by the principal and School Improvement Team.
- Academic grades shall be a content-based measure of what students are able to demonstrate.
- Grading shall provide appropriate and accurate feedback to the students and parents.
- Grades shall reflect performance that is consistent with demonstrated achievement on the goals of the North Carolina Standard Course of Study.

Table 10

Ken O'Connor's Evaluation of the District's Reteach/Retest Policy

Overall Grading guideline expectation	Comments from O'Connor
Grading practices shall be reviewed annually by the principal and School Improvement Team.	"An annual review is a good idea"
Academic grades shall be a content-based measure of what students are able to demonstrate.	"I would have thought that that should've said standards-based measures, given the basis for schooling now."
Grading shall provide appropriate and accurate feedback to the students and parents.	"I agree completely but the way we do that is first of all with words not symbols, not points or percentages and if we are going to have any sort of labeling ... then that should be related to levels of proficiency because that gives real information, if we are talking about appropriate and accurate feedback ... its helpful to know- are you proficient or not ... it's not very helpful to know that you have 78% or 76%."
Grades shall reflect performance that is consistent with demonstrated achievement on the goals of the North Carolina Standard Course of Study.	"Good ... to me that is sort of saying the same things as the second one but just bringing it to the state."

Reflections from Participants

Toward the end of each interview, participants were asked to reflect on their grading practices and share what aspect of their grading they would change if they could. Anders's response was deeply reflective and, in a sense, seemed to express some level of regret and sadness. He shared,

I had a student who had the highest average in my class but didn't do very well on the EOG. She did all of her work and worked really hard. She just (pause) when it came time for the EOG she just didn't (pause) to me that stands out, you know, did I fail her? Did I not give her the attention she needed just because she's like doing everything she's supposed to do? She's turning everything in. Maybe I didn't read close enough her responses on short answers. I think for her ... her grade reflected effort more than it did mastery.

Anders went on to explain that he would like to make changes that would be more aligned to state standards.

According to Hanna, when asked about the changes she would make to her grading, she had this to say: "I'd get rid of tests!" When asked to explain this in more detail, she appeared to be suggesting that she would assess in more ways other than traditional paper and pencil tests. She went on to explain that she would utilize projects more to keep students engaged and give them more creative ways of demonstrating their learning:

I think that you kill a student's curiosity by constantly throwing questions at them like that. I think that's the way that it has always been, and I find that particularly with science; little kids come in with an amazing curiosity ... why? Why is the

sky blue? ... and as they get older that curiosity goes away.

She went on further to explain, “When students are allowed to explore in ways that they find more engaging, they learn more and are more motivated to do the work. When we ask the questions of students on a test, they tend to become disengaged.”

Parker expressed that she would want her grades to be more aligned to the standards she teaches. All other teachers commented on homework being a struggle to manage and expressed that they would like to see this eliminated.

Axial Coding

According to Charmaz (2014), this is a method of coding that “treats a category as an axis around which the analyst delineates relationships and specifies the dimensions of this category” (p. 341). During the axial coding stage, one major category, policies, was selected as the core category or central phenomenon. Corbin and Strauss (2015) explained that the core category is the “concept that is abstract and broad enough to be representative of all participants in the study” (p. 188). Furthermore, this category should “have the greatest explanatory power and ability to link other categories to it and to each other” (Corbin & Strauss, 2015, p. 188).

After selecting the core category, the researcher then revisited the data to establish links and to identify additional categories related to this central phenomenon. Through axial coding, data were integrated “into a more coherent whole after the researcher has fractured them through line by line coding. Axial coding is a procedure applied to the data rather than emerging from the data” (Charmaz, 2014, p. 341).

Axial coding categories, according to Creswell and Poth (2018), “relate to and surround the core phenomenon in a visual model called the axial coding paradigm” (p.

85). Figure 2 depicts the axial coding diagram.

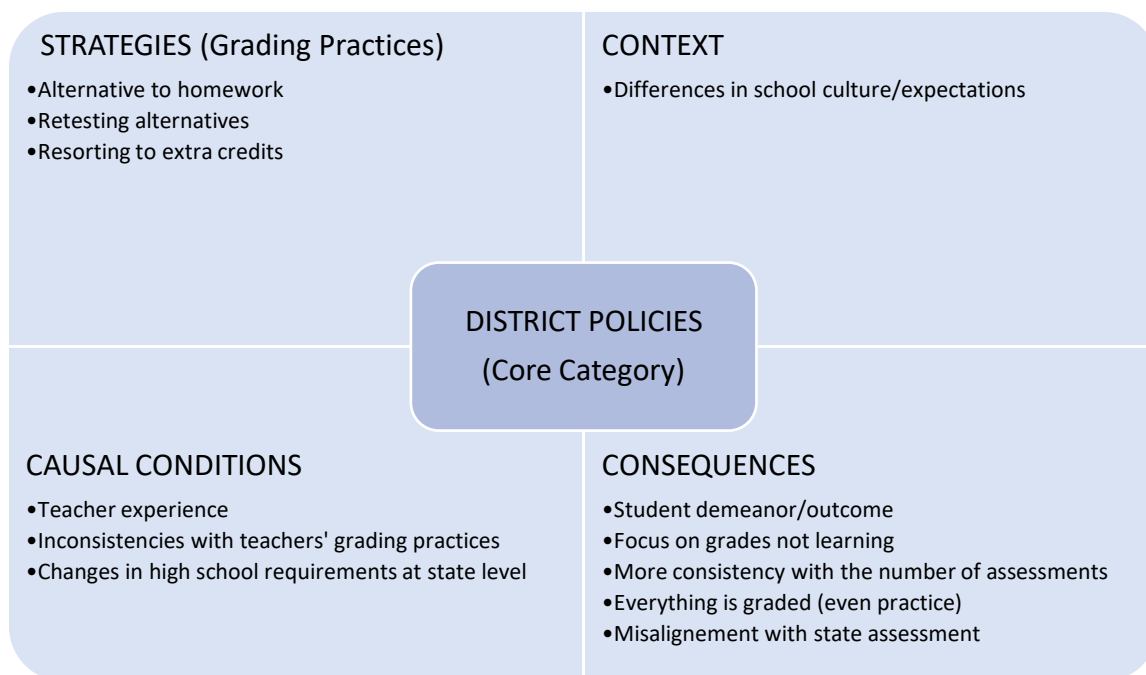


Figure 2. Axial Coding Paradigm.

Central phenomenon. The axial coding paradigm depicts district policies as the central phenomenon/core category and the other related subcategories. This was selected because of its far-reaching effect on teacher grading practices. All habits, strategies, and actions taken by teachers and administrators can be traced back to the overarching district policies.

Strategies. The diagram shows that the district policies influence and dictate teacher grading practices by providing the parameters within which teachers must evaluate student performance. In response to requirements set forth in the district policies, teachers employ various strategies when they grade. The policies require that grades include scores from homework even though many students frequently fail to complete these. In response to this issue, teachers strategically provide time and space

for homework to be done in the classroom. Teachers in Title 1 schools have explained that there are no guarantees that students will have the means and ideal settings needed to get this done at home; hence, from an equity standpoint, teachers make adjustments to meet the needs of their diverse student population.

Context and causes for the policies. During an interview with one district leader, it was explained that these policies had to be put into place in order to reduce the variations in teacher grading practices. Previously, when no limit was placed on the number of grades needed, it was possible that some teachers could report grades that included one test grade, while others could include four or five. It was also possible that students could fail a semester due to student failure to complete homework because homework had the same weighting as all other categories. These practices led to the establishment of the current district policies.

Consequences. Some of the consequences of current district policies include more consistency with the number of assessments, more grades are now required which has led to a greater focus on grades rather than learning, and fewer opportunities for practice without penalties.

Propositions

In grounded theory studies, selective coding is used to make the leap from conceptual categories to a theoretical explanation of the process (Corbin & Strauss, 2015). The selective coding analysis of the data resulted in the following hypotheses or propositions to explain the influences on teacher grading practices.

1. Teachers must operate within the confines of the district's grading policies.

The policies therefore exert the greatest influence on teacher practices and

outweigh other influential factors such as professional development, colleagues, or personal experiences as a student.

2. The district grading policy places greater emphasis on tests, which represent 40% of a student's grade, compared to other assessments. Teachers also regard scores from tests as being more accurate than scores from classwork, guided work, homework, or group work; yet with the policies that are in place, teachers are compelled to grade other assignments that do not necessarily reflect student achievement. In this grading system, when teachers work toward meeting the requirements of the district policies, they naturally develop the compulsive habit of grading everything. This then leads to greater student apathy and students not wanting to turn in work out of fear that their practice will become a summative judgement of their achievement and lead to permanently low scores. To rectify this issue, teachers therefore succumb to the practice of giving extra credits for work habits in the hopes that this will help increase student motivation.
3. Even though grading practices vary among teachers from different grade levels and subject areas, one constant that emerged from interviews was the inclusion of effort and other nonachievement factors in the grading process. Core teachers in Title 1 schools grapple with low student completion rates on homework. Teachers are aware of inequitable conditions outside of school that contribute to this; despite this knowledge, they are compelled by district policies to grade homework. In response to this, teachers resort to informal approaches to grading, which includes the consideration of student behaviors

when determining student grades.

Selective Coding Visual Model

The model is purposely composed of triangles to symbolize the call for change. The top of the model depicts the central phenomenon and major influence on teacher grading practices/policies. The central phenomenon, policies, was placed at the top of the diagram to symbolize that the greatest influence on teacher grading practices starts from the top with school and district leadership.

The model shows that the policies dictate and influence the grading practices that teachers employ and simultaneously produce various student outcomes and problems which are categorized as predicaments in grading. These were also outlined by O'Connor and teachers during interviews.

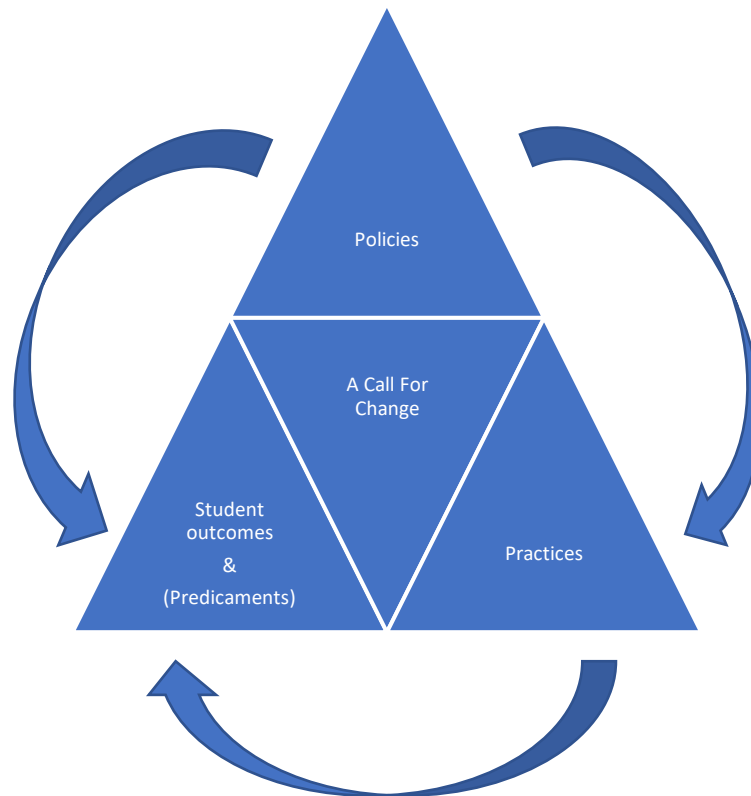


Figure 3. Visual Model Showing the Influences on Teacher Grading Practices.

Summary of Findings

This chapter presented the findings of this study with respect to the three research questions and levels of grounded theory analysis.

Through open coding analysis, three major categories (policies, practices, and predicaments) arose that are grounded in the data from participants as evidenced by direct quotes from interviews and documents.

The axial coding analysis was used to show the links between and around the central phenomenon and was used to develop the theory. The axial coding categories were then selectively coded to identify the most significant phenomenon which influences the development of teacher grading practices.

Chapter 5: Conclusions

This chapter presents a summary of the major findings from the study, a reexamination of the research questions which guided the study, theoretical connections, and concluding remarks. The chapter includes an introduction, research findings and theoretical implications, and recommendations and ideas for future research.

Introduction

In previous chapters, the literature on grading was presented, thereby making the case that a problem exists with traditional grading methods. In addition, it has been shown that there was a need for conducting a study that not only describes teacher grading practices but also explains the influences on those practices.

Concerns about the reliability of grades in communicating meaningful information about student achievement have been raised for a long time (Starch & Elliot, 1912, 1913). Additionally, several references have been made in the literature over the years about the prevalence of traditional grading practices and the predicaments they present: Nearly 2 decades ago, Marzano (2000) highlighted that our current “grading system is at least 100 years old” (p. 13) and that it represents a part of the “educational history that has been almost impervious to change” (Farr, 2000, p. 3). Additionally, Ohlsen (2007) has referred to traditional classroom grading practices as being “dismal for decades” (p. 5). More recently, Reeves et al. (2017) regarded traditional grading as “the wild west of school improvement” (p. 44).

The researcher carried out this study to determine if traditional grading practices were in fact the predominant grading method being used by Title I teachers as the literature suggests. If the findings were consistent with the literature, the researcher

wanted to be able to explain why teachers continue to embrace traditional grading methods.

The primary purpose of this qualitative study was to describe Title 1 teacher current classroom grading practices and explain the process by which their practices develop. A secondary purpose of the study was to describe and analyze the existing district grading policies to determine the extent to which these policies influence and perpetuate traditional grading practices among teachers in Title 1 schools.

Through grounded theory methods, a theoretical model was developed and presented to explain the influences on teacher grading practices. The three research questions which guided the study were

1. What are the grading practices currently used by teachers?
2. What factors influence the development and evolution of teacher grading practices?
3. What changes (if any) might be implemented in current grading practices which would provide a better measure of assessing student achievement?

The findings from this study are presented in terms of these research questions.

Summary of Findings

Research Question 1: What are the grading practices currently used by teachers? The findings from the study were consistent with the literature which suggests that teacher grading practices remain largely unchanged and traditional.

Formative and summative assessments. During a 9-week grading period, teachers go through a cycle of planning, teaching, grading, and reporting. Teachers are required to report grades each week through an online grade book so students and parents

are able to access grades. Teachers have shared that they value grades from tests as the best indicators of student achievement, compared to grades from other sources like classwork, homework, or group work. Tests account for 40% of a student's overall grade; however, because numerous grades from other sources of assessments are also required, teachers report that they frequently feel rushed to satisfy the minimum requirement of three test grades per 9-week grading period. In addition, leading expert and author Ken O'Connor has spoken out against the number of grades teachers in this district are required to report during each 9-week period. His expert opinion is that requiring so many grades leaves little time for students to engage in a proper learning cycle which includes practice (not be graded), feedback from the teacher (also with no grades attached), more practice, assessment, reteaching, and reassessment if needed. This implies that tests, which possibly provide the best measure of student achievement relative to homework, classwork, and other forms of practice are not always administered under ideal conditions; that is, if teachers feel rushed to simply satisfy the minimum required number of assessments, this leaves very little opportunity for timely feedback, additional practice, and reassessment. This possibly explains why, in an attempt to resolve this challenge, teachers have had to adapt certain grading practices and make concessions such as offering test correction opportunities or allowing students to retake the same test, all in lieu of conducting proper reteaching and retesting exercises.

In summary, the district requires a minimum of 12 assessments in addition to homework. The data from the study suggest that requiring this amount of assessments (all of which may not be achievement grades), together with teacher compulsions to grade everything, only dilutes the quality of the overall grade, compromises the teaching

learning process, and therefore renders the grade less meaningful.

Second chances and the most recent evidence. According to Guskey (2015), “conversations with teachers about grading typically reveal a staggeringly diverse array of policies and practices” (p. 97). The researcher found this to be especially true as it relates to teacher individual policies concerning reassessments and the manner in which they are graded.

Several teachers reported that instead of reteaching and then reassessing when students fail to perform well on a test, they allow students to do test corrections. The manner in which they handle grading this alternative reassessment varies greatly from one teacher to the next. Many teachers reported that students were given the average of the two grades (the first test grade and the reassessment grade); some reported they placed a limit on the highest grade the student could receive after a reassessment (one teacher gave students no more than 60% after the reassessment); other teachers reported they gave students half the points back for every question that was corrected. When teachers were asked to explain their reassessment practice, they all stated they did not think it was fair to other students who mastered it on the first attempt. Teachers also reported they did not want students to become overly reliant on the reassessment opportunities.

This idea was also expressed by one district leader when she explained why the district policy placed a limit on the highest score a student could earn on a retest in high school. There is no district policy to guide teachers on how reassessments should be carried out at the middle school level, but a policy for retesting exists for high school. One aspect of this policy stipulates that the “highest grade that a student may make on a

re-test is 75%.” During the interview with Dr. Bentley, she explained this was put into place to ensure that “children give their very best the very first time and that it’s a true reflection of their ability.” In addition, she emphasized that if research exists that suggests there is a better cut-off grade, the district would be willing to adjust the policy as long as it does not “damage the child’s GPA, they can still recover, but they’re not getting an advantage over kids who took it the first time.” According to Dr. Bentley, “While we want them to have a second chance, we don’t want them to get accustomed to second chances because, is that really the real world if they go to college or to a career?” Contrastingly, experts have explained that in world outside the classroom, people frequently get second chances. Vatterott (2015) argued, “Drivers are not restricted by how many times they took the driver’s test and their scores are not averaged together” (pp. 70-71). Additionally, “A lawyer’s license to practice law does not indicate how many times he or she took the bar exam” (Vatterott, 2015, p. 71). Why then do we not treat school reassessments in a similar manner?

Prominent individuals in the field of education continue to argue that the most recent evidence of learning is the most accurate and that grades from reassessments should replace old evidence (Guskey, 2015; O’Connor, 2011; Reeves, 2011; Vatterott, 2015). Similarly, Reeves et al. (2017) contended that “a grading system that emphasizes a student’s current performance or most recent evidence ... gives students a reason to keep trying” (p. 43); however, one “that persistently punishes mistakes instead of rewarding eventual progress and mastery guarantees the stagnation of learning (Reeves et al., 2017, p. 43).

The researcher strongly believes that teachers and school leaders attempt to

develop grading policies that reflect student learning. There was no indication to suggest that teachers/leaders have purposely established schemes and policies to intentionally hurt students; however, practices like these can be harmful to students and prove counterproductive to learning.

Inclusion of nonachievement factors in grades. Many teachers reported that they include nonachievement factors in the grades they report. This is especially evident in statements from Hanna and John. According to Hanna, “If a student gives me some kind of effort ... I will give you some credit for the effort because I don’t want them to feel completely defeated.” Similarly, John explained, “If you attempted it, and you put something down and you show me you’re using your time efficiently to do that, I will give you a 50 even if you get it all wrong.”

According to O’Connor (2011), teachers combine achievement and nonachievement factors into grades for various reasons: “One belief is that this practice appropriately rewards students who are well behaved and punishes those who do not behave as expected” (p. 17). This belief is reflected in the following statement by Anders:

When I see that she’s working really hard or he didn’t talk the whole class, but they just didn’t do this correctly, I don’t want them to fail ... so the kid who worked really hard, I always give them a better score.

A second reason O’Connor (2011) offered is that “teachers have had no way to communicate separately about the behaviors they think are important and so have blended them together with achievement” (p. 17). Various documents including teacher grade books, syllabi, and the district policies confirm that teachers in the district report

student grades on a traditional report card which shows the average of all types of scores combined. Data from these documents revealed that teachers had no alternative way of reporting behaviors and other nonachievement factors separately.

Zeros and missing work. Teachers in the study reported that the grades assigned to students not only represent the average of various types of assignments but also frequently included zeros on a percentage grading scale. A typical statement that confirms this behavior was provided by Hanna when she explained conditions under which students are assigned a zero: “A student will get a zero ... if you just don’t do the assignment at all, you make no effort whatsoever ... you are getting a zero. Period.” The majority of the teachers interviewed also reported that students rarely complete homework assignments and therefore frequently earn a zero for missing homework. Not only do teachers assign a zero, but no policy exists at the district level that prohibits it. This is evident in a statement from Dr. Bentley regarding the conditions under which a grade of zero could be assigned: “If a child chooses not to make something up and they’ve been given ample opportunity, a teacher could give a zero.”

O’Connor (2011) reported that there are three problems with zeros: “Zeros give a numerical value to something that has never been assessed ...; they can have counterproductive effects on student motivation, and they involve inappropriate mathematics” (p. 96).

Other experts have spoken out against the practice of assigning zeros. Guskey (2015) explained that grades should communicate information about student achievement. Grades should not punish students, but “in a percentage grading system, assigning a grade of 0 does exactly that” (p. 32). Furthermore Wormeli (2006b) attested

that a zero can have a devastating impact on students: “No matter what the student does, the grade distorts the final grade as a true indicator of mastery” (pp. 137-138).

Even though teachers reported that they assign zeros when evidence is missing or work is not turned in, they also acknowledge that this might lead to student apathy and decrease motivation. Teachers have also reported that they reward students with extra credits to reinforce desirable behaviors. They assign extra credits to motivate students and to encourage them to keep trying; however, when students have become disengaged or do not make sufficient effort, they are punished with zeros. The true achievement grade becomes lost somewhere in between the two extremes as depicted by Figure 4.



Figure 4. Distortion of Achievement Grade.

These findings seem to suggest that teacher grading practices, policies, and habits have become deeply steeped in long-held grading traditions, to the extent that teachers are not always acting with complete awareness of the consequences of their actions.

Research Question 2: What factors influence the development and evolution of teacher grading practices?

District policies. Findings from the study suggest that while there are numerous influences on teacher grading practices, the main factor which influences teacher grading practices is the district’s grading policies/guidelines. Teachers are compelled to work

within the confines and parameters set forth in the district grading policies.

The district policies require teachers to collect evidence through multiple assessment methods: tests, quizzes, group work, teacher-led activities, daily classwork, projects, common assessments, and homework. The scores from these assessments are then averaged together to produce a final grade as depicted in Figure 5. Even if a teacher held the belief that daily classwork, homework, and other forms of practice should be formative in nature and therefore not be included in a student's overall grade, they would still be compelled to consider and include these sources of evidence since they are required by the district policies.

This practice is consistent with the literature from Guskey (2015), which stated, "Teachers today draw from many different sources of evidence in determining students' grades" (p. 72). Other experts attest that this leads to a hodgepodge grade which includes elements of achievement and nonachievement factors (Brookhart & Nitko, 2008; Cross & Frary, 1999; Guskey, 2011). Similarly, Vatterott (2015) explained that "when teachers grade everything, the grade means nothing" (p. 34). Further, when everything is graded, it allows some students to "manipulate the system and mask poor academic performance" (Vatterott, 2015, p. 34). This implies that in this system, students could earn high grades for simply turning in all their work and being well behaved (O'Connor, 2011; Vatterott, 2015). Sadly, this was also reflected in the following statement from Anders:

I had a student who had the highest average in my class but didn't do very well on the EOG. She worked really hard.... Did I not give her the attention she needed just because she's ... doing everything she's supposed to do? She's turning everything in ... I think for her ... her grade reflected effort more than it did

mastery.

Figure 5 depicts the overall grade reported by teachers. The grade produced represents a confluence of achievement factors and nonachievement factors which is consistent with the requirements of the district grading policy.

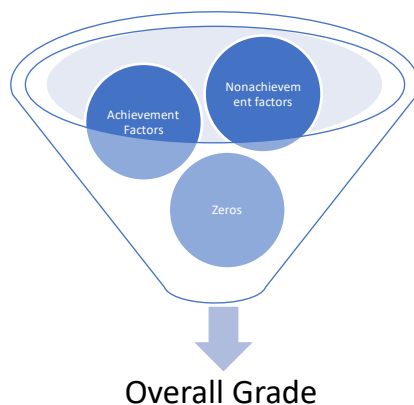


Figure 5. The Overall Grade – A Confluence of Factors.

The absence of discussions – let’s talk about grades. Guskey (2015) pointed out that “school leaders are generally reluctant to question teachers on the inclusion of nonachievement factors in academic grades” (p. 97) and “most consider such policies to be part of well-established grading traditions” (p. 97).

One expectation expressed in the district grading policy is that grading practices should be examined annually by the principal and school improvement team. Notwithstanding, findings from the study revealed that teachers seldom engage in discussions about their grading practices.

The researcher strongly believes that the absence of discussions about grading has greatly influenced teacher grading practices and contributed to the perpetuation of outdated and damaging habits.

By the end of the study, the researcher gained a better understanding of teacher grading practices and factors influencing these: Teacher grading practices are guided by district policies which require that all grades be averaged on the percentage grading scale; the policies also encourage teachers to grade everything (homework, classwork, group work, quizzes, lab work, practices, projects, and tests); and weaknesses and gaps in the district policies become replaced with unique or traditional practices passed down from teacher to teacher. In addition to the knowledge gained about teacher grading practices, questions have been provoked in the mind of the researcher. Could teacher grading practices improve through discussions about the purpose of grades? Could teacher grading practices improve if provided with literature on best grading practices? Would schools become a place where students are free to be more focused on learning and less worried or concerned about grades and being penalized for not knowing it on the first attempt? What would be the result, if teachers provided nonpunitive timely feedback on practices to help guide students toward mastery of a learning target, rather than grade students while the learning is underway?

Some of these questions have also helped to shape the researcher's recommendations for future research.

Research Question 3: What changes (if any) might be implemented in current grading practices which would provide a better measure of assessing student achievement?

Grading policies. Data provided during an interview with grading expert Ken O'Connor strongly suggested that making changes to the district grading policy and engaging in discussions with teachers about the purpose of grades could produce a major

shift in teacher grading practices that better support learning. This sentiment is also shared by McElligott and Brookhart (2009, as cited in Guskey, 2009), who stated that “Schools and districts should have grading policies and review them regularly” (p. 67). Furthermore, McElligott and Brookhart (2009, as cited in Guskey, 2009) contended that these policies “should be of substance, treating such things as what should go into a grade (achievement measures, not behavior for example)” (p. 67).

Currently, the district grading policies compel teachers to include nonachievement factors such as daily classwork practice, group work, and homework in the grades reported. This practice is inconsistent with recommendations from grading experts who explain that when nonachievement factors are included in the overall grade, it obscures the meaning of the grade and compromises their communicative value. According to Tomlinson and McTighe (2006),

A grade should give as clear a measure as possible of the best a student can do.

Too often, grades reflect an unknown mixture of multiple factors.... The problem transcends individual teachers. Unless teachers throughout a school or district completely agree on the elements and factors them into their grading in consistent ways, the meaning of grades will vary from classroom to classroom, school to school. (p. 133)

To gain a better understanding of changes that need to be made in order for grades to be a better measure of student achievement, it is necessary to outline the current problems.

The district policies compel teachers to use percentage grades. All teachers have reported that they use a percentage grading scale when reporting on student achievement.

According to Guskey (2015), “Percentage grades give the illusion of precision” (p. 23) and “represent ... the most significant deterrent to the development of better, more honest, and more reliable grading” (p. 24). Instead, experts recommend a return to past practices. According to Reeves et al. (2017), “The more appropriate ... way to solve the problem is to return to the time-honored grade-system in which A is 4, B is 3, C is 2, D is 1 and F is 0” (p. 44).

The district policies compel teachers to consider and collect at least 12 pieces of evidence during each 9-week grading period and average them all together to produce an overall grade (see Figure 5). With respect to the number of assessments required by teachers, grading expert Ken O’Connor stated during the interview that this is “the single worst thing” in the district policies. Guskey (2015) explained, using an analogy, why combining so many different evidences produce misleading information about student achievement: “If someone proposed combining measures of height, weight, diet, and exercise into a single number or mark to represent a person’s physical condition, we would consider it laughable” (p. 74); yet despite this knowledge, the district policies require teachers to do just that. Therefore, it would be far better if the policies required fewer pieces of summative evidence that are linked to standards and which only include achievement grades.

Tradition, beliefs about grading, and mindset. Much has been said in this study about the impact that policies have on teacher grading practices. Notwithstanding, teachers and administrators still have the latitude and many means by which improvements can be made.

The district policy does not specify that teachers should consider effort in

determining grades, yet teachers do; the policy does not specify that teachers should reward students with extra credits for behavior and other factors unrelated to achievement, but teachers do; and the policy does not specify that zeros should be given as place holders in a grade book for missing work, yet teachers report that this is commonly practiced.

Many of these traditional habits embraced by teachers are not explicit requirements under district grading policies. Even though district policies provide parameters that influence how teachers grade, there are traditional habits teachers have embraced that cannot be attributed to any policy requirements. Accordingly, there is room for improvement even within the confines of those district policies. Teachers have the power to exclude effort, extra credits, and zeros from grades. These are changes that can be made with little effort, if teachers and leaders engage in discussions about the purpose of grades and become more knowledgeable of the impact that these habits can have on learning. Eliminating these practices could be simple fixes at the classroom level that would contribute to making grades more meaningful and accurate.

Additionally, experts suggest that when district leaders, principals, and teachers take steps to clarify and define a common purpose for grading, practices will start to improve and be more aligned with those purposes. When everyone embraces the belief that the purpose of grading is to reflect academic achievement, teachers will no longer be required to consider and include classwork, homework, and other forms of nonachievement factors in grades. Teachers will no longer grade practice but instead use these formatively and for feedback (Vatterott, 2015).

Implications

The data gathered from this study indicate that the grading practices employed by Title I teachers are predominantly traditional, as the literature suggested. The traditional grading methods referenced in the literature and that were being employed by teachers in the study included the use of the arithmetic mean, or average, to calculate a final grade; the grading of practice (homework, classwork, group work, formative assessments); the use of the zero on the 100-point scale; and the inclusion of nonachievement factors in grades. This finding implies that the grading practices of teachers in this district, if unchanged, will continue to jeopardize the reliability of grades and therefore “weaken the link between grades and academic achievement” (Welsh, D’Agostino, & Kaniskan, 2013, p. 27).

Further findings also indicated that a very powerful influence on teacher grading practices is the district grading policies. These guidelines/policies dictate how teachers report grades and the composition of those grades. The current district policies not only influence but ensure that teachers continue to use some traditional grading practices. Experts recommend that if we want grades to support learning, only summative evidence should be included in a student’s reported grade. All other forms of assessments (formative, practice, and homework) should be excluded from a student’s final grade. If grades are to be more meaningful, they should be organized by standards not assessment type. This recommendation is consistent with the literature by O’Connor (2011) and Guskey (2011).

Contrastingly, the district grading policies specify that teachers are expected to grade practice assignments like formative assessments, homework, and classwork. The

grading policy also specifies that scores from all these types of assessments should be averaged to produce a final grade. If current district policies remain unchanged, traditional grading habits will continue to prevail.

Recommendations

The following recommendations for practice center around one major underlying theme in this study: a call for change.

If teacher grading practices are in need of improvement, one recommendation is that the policies and guidelines for grading in the district be altered to reflect the recommendations of grading experts. This recommendation is consistent with the literature by Guskey (2011): “Even in schools where established policies offer guidelines for grading, significant variation remains” (p. 53) in teacher grading practices. Furthermore, he contended that in some cases the policies are inherently flawed: “Rarely do these policies and practices reflect those recommended by researchers” Guskey (2011, p. 53). If a district wants to encourage better grading practices among teachers, the grading policy guiding their actions should require it:

To succeed in tearing down old traditions, you must have new traditions to take their place. This means that education leaders must be familiar with the research on grading and what works best for students, so they can propose more meaningful policies and practices that support learning. (Guskey, 2011, p. 21)

A change in policy should also accompany a change in mindset among teachers, administrators, and district leaders as it relates to best grading practices. One finding from the study revealed that despite being one of the requirements listed in the district policy, conversations about grading practices rarely occur at the school level with

administrators and teachers. Consequently, schools currently operate without a clear or common purpose for grades.

It is recommended that administrators and teacher leaders take steps to initiate grading conversations in order to affect change within their schools and develop grading practices that are more reflective of student achievement. This could start with the establishment of committees at individual schools comprised of teacher leaders and administrators who engage in reading current research about grading practices and discussions about the purpose of grades. Engaging in more frequent and focused discussions can help teachers and administrators to reflect on and expose conflicting beliefs about grades. When school administrators examine the grading practices within their own building, they will gain a more comprehensive awareness of how student grades are determined.

As O'Connor (2011) pointed out, grades have traditionally served multiple purposes; and Brookhart (2004) also explained that "it is very difficult for one measure to serve different purposes equally well" (p. 21). When teachers and administrators agree on the purpose of grades, the appropriate grading methods usually follow that purpose (Brookhart, 2004; Guskey, 2015; Vatterott, 2015). Many attempts have been made to help teachers understand the purpose of grades. This has been addressed repeatedly in the literature (Brookhart, 1993; Cross & Frary, 1999 1996; Guskey, 1996; Marzano, 2000; Stiggins, 2001); however, this study reveals that very little progress has been made in actual classrooms grading practice.

Recommendations for Future Research

The body of research on grading continues to grow, but still more work remains.

The following represents recommendations for future research which emerged from the findings of this study:

1. What effect would professional learning and development have on teacher grading practice? Data from the study, in part, suggest that some of the practices teachers employ stem from a lack of knowledge about best grading practices. Further research is needed to determine the change in grading habits, if any, that would occur if teachers were to participate in a course of study that focuses on current research pertaining to grading. If teachers were more knowledgeable about the literature on grading, would this lead to a change in their practice? More research is needed to help explore this question.

A mixed-methods study could be conducted by first administering a survey that would provide an overview of teacher grading practices. This result could then be sorted based on grading practices, and groups of teachers could be selected to undergo various forms of professional development. At the end of the course or treatment, teachers could share reflections about the changes in their grading practices they have implemented, if any. Data from this type of study could inform professional development activities for both new and experienced teachers. This proposed study could involve the following:

- a. Administer a pretest and posttest in the form of a survey to gather information about teacher levels of knowledge about best grading practices before the course of study and after. The survey could also

gather data to determine the grading practices teachers currently use.

- b. Design a short course of study which includes a selection of current articles, books on grading, and reflection questions to guide participant learning. Participants would not only complete readings, but they would also maintain a journal to document their beliefs, thoughts about grading, and how they change, if any, throughout the duration of the course. In addition, participants would be required to document and engage in discussions with other participants in the group. These discussions would provide useful data and therefore could be documented.
 - c. At the end of the study period, teachers would complete a survey to help measure changes in their level of knowledge about best grading practices. In addition, teachers would reflect on changes in their beliefs, if any, and report about the grading practices they would be willing to change.
2. When teachers were asked about the grading challenges they experienced, a common concern expressed was student apathy. Hanna's response was typical of the sample of teachers interviewed when she shared that one of her major struggles is "just getting kids to do their work." Based on this shared concern among teachers, it would be beneficial to carry out a similar study that includes students. The inclusion of students in the study would not only help to address the concern of student apathy but would also provide more comprehensive data on the topic. This perceived disinterest/disengagement among students has prompted some teachers to give extra credit points in hopes that this would alleviate the problem and motivate students to work.

For other teachers, the consequence meted out to students included assigning zeros. It would be enlightening if data were available that could provide a better understanding of student perspectives on this issue. How do students experience grading? What changes to teacher grading practices would students recommend that would better support their learning? If students understood that their teachers' main concern is to help them be more successful, would this spark a change in student behavior?

Several teachers shared that they disliked giving homework because students do not do it, which leads to them getting a grade of zero for this missing work. This in turn compromises student-teacher relationships in some cases. It would be very helpful to know if students value homework and if they would be more inclined to doing it if there was no grade associated with it. If homework was treated as practice and not graded, would students be more willing to complete it with fidelity? Teachers, like Anders, have shared that students frequently ask, "Is this being graded?"; and if the answer is no, students rarely complete the task. Common concerns like these could be further explored and better understood if students were to be included in the study. Therefore, it is recommended that a similar grounded theory study be conducted to describe and explain teacher grading practices and how their students experience and respond to those methods.

3. This study needs to be replicated with a larger sample of teachers, including teachers from schools that are not classified as Title 1 schools. This would allow the researcher to better examine differences in grading practices and

establish which practices are more prevalent among teachers. It would be worthwhile to explore the types of concessions made, if any, for students in Title 1 schools compared to students in other schools. Participants in the study have expressed that they consider nonachievement factors in their grading, as a way of helping students to recover from failing grades. What other concessions are being made for students? Is this the same at other schools with a larger percentage of high-achieving students? Would the theoretical propositions from this study be applicable to other school populations?

4. The study was limited to teachers of core subjects such as mathematics, language arts, social studies, and science. Further research is needed to compare the grading practices of core subject teachers with teachers in other disciplines, such as music, art, and physical education. Could their grading practices be vastly different from teachers in other disciplines?
5. The study has confirmed that traditional grading practices are the predominant grading methods being used by Title 1 teachers. There is an urgent need for change. Reeves et al. (2017) asserted that this change “will benefit all learners, but especially those who are struggling.” (p. 45). Given the potential benefits and the opportunities which exist in Title 1 schools, more research on grading is needed with this population of teachers and students. One next step would be to design a study which involves a group/class of students from a Title I school who will experience traditional grading during the first half of the year and nontraditional grading during the second half of the school year.

This could potentially serve as a form of a Title I reform effort. The nontraditional methods would involve eliminating grades from sources like homework, classwork, group work, and other forms of practice. In this system, teachers would continue to offer feedback to students to help them improve but no grades would be associated with this feedback or any work that is for practice. The practice of using zeros on the percentage grading scale and extra credits would also be eliminated during the nontraditional treatment. The only grades that would be reported in this system would come from summative assessments directly related to standards taught. In the nontraditional grading system, no more than four or five pieces of evidence would be required per 9-week grading period. Surveys could be conducted with students and teachers to assess the effectiveness of this grading method on student achievement and their level of motivation.

References

- Bailey, M. (2012). *The relationship between secondary school teacher perceptions of grading practices and secondary school teacher perceptions of student motivation* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3504337)
- Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80(2), 139-148.
- Bowers, A. J., Sprott, R., & Taff, S. A. (2013). Do we know who will drop out? A review of the predictors of dropping out of high school: Precision, sensitivity, and specificity. *The High School Journal*, 96(2), 77-100.
- Brimi, H. M. (2011). Reliability of grading high school work in English. *Practical Assessment, Research and Evaluation*, 16(17), 1-12.
- Brookhart, S. M. (1991). Grading practices and validity. *Educational Measurement: Issues and Practice*, 10(1), 35-36.
- Brookhart, S. M. (1993). Teachers' grading practices: Meaning and values. *Journal of Educational Measurement: Issues and Practice*, 30(2), 123-142.
- Brookhart, S. M. (1994). Teachers' grading: Practice and theory. *Applied Measurement in Education*, 7(4), 279-301.
- Brookhart, S. (2004). *Grading*. Upper Saddle River, NJ: Pearson/Merrill/Prentice Hall.
- Brookhart, S. M. (2011). Starting the conversation about grading. *Educational Leadership*, 69(3), 10-14.
- Brookhart, S. M. (2015). Graded achievement, tested achievement, and validity, educational assessment. *Educational Assessment*, 20(4), 268-296.

- Brookhart, S. M., Guskey, T. R., Bowers, A. J., McMillan, J. H., Smith, J. K., Smith, L. F., ... Welsh, M. E. (2016). a century of grading research: meaning and value in the most common educational measure. *Review of Educational Research*. 86(4), 808-848.
- Brookhart, S. M., & Nitko, A. J. (2008). *Assessment and grading in classrooms*. Upper Saddle River, NJ: Pearson Merrill Prentice Hall.
- Charmaz, K. (2014). *Constructing grounded theory*. Los Angeles: Sage.
- Cizek, G. J. (1996). Grades: The final frontier in assessment reform. *NASSP Bulletin*, 80(584), 103-110.
- Clymer, J. B., & Wiliam, D. (2007). Improving the way we grade science. *Educational Leadership*, 64, 36-42.
- Corbin, J. M., & Strauss, A. L. (2015). *Basics of qualitative research: techniques and procedures for developing grounded theory*. Los Angeles: Sage.
- Cox, K. B. (2011). Putting classroom grading on the table: A reform in progress. *American Secondary Education*, 67-87.
- Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. Boston, MA: Pearson.
- Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry & research design: Choosing among five approaches*. London: Sage.
- Cross, L. H., & Frary, R. B. (1999). Hodgepodge grading: Endorsed by students and teachers alike. *Applied Measurement in Education*, 12(1), 53-72.
- Dueck, M. (2014). *Grading smarter, not harder: Assessment strategies that motivate kids and help them learn*. Alexandria, VA: ASCD.

- Durm, M. W. (1993). An A is not an A is not an A: A history of grading. *The Educational Forum*, 57, 294-297.
- Dyb, C.D. (2011). *Teacher perceptions of middle level grading practices*. (Doctoral dissertation, Edgewood College). Retrieved from ProQuest Dissertation and Thesis. (UMI No. 3498285)
- Farr, B. (2000). Grading practices: An overview of the issues. In E. Trumbull & B. Farr (Eds.), *Grading and reporting student progress in an age of standards* (pp. 1-21). Norwood, MA: Christopher-Gordon Publishers.
- Finkelstein, I. E. (1913). *The marking system in theory and practice*. Baltimore: Warwick & York.
- Frisbie, D. A., & Waltman, K. K. (1992). Developing a personal grading plan. *Educational Measurement: Issues and Practices*, 11(3), 35-42.
- Goodwin, B. (2011). Grade inflation: Killing with kindness? *Educational Leadership*, 69(3), 80-81.
- Guskey, T. R. (1994). Making the grade: What benefits students? *Educational Leadership*, 52(2), 14-20.
- Guskey, T. R. (1996). *Communicating student learning*. ASCD yearbook. Alexandria: VA: Association for Supervision and Curriculum Development.
- Guskey, T. R. (2000). Grading policies that work against standards and how to fix them. *NASSP Bulletin*. 84(620), 20-29.
- Guskey, T. R. (2001). Helping standards make the grade. *Educational Leadership*, 59(1), 20-27.

- Guskey, T. R. (2004). 0 alternatives. *Principal Leadership: High School Edition*, 5(2), 49-53.
- Guskey, T. R. (2006). Making high school grades meaningful. *Phi Delta Kappan*, 87(9), 670-675.
- Guskey, T. R. (2009). Grading policies that work against standards and how to fix them. In T. R. Guskey (Ed.), *Practical solutions for serious problems in standards-based grading* (pp. 9-26). Thousand Oaks, CA: Corwin.
- Guskey, T. R. (2011). Five obstacles to grading reform. *Association for Supervision and Curriculum Development*, 69(3), 16-21.
- Guskey, T. R. (2013). The case against percentage grades. *Educational Leadership*, 71(1), 68-72.
- Guskey, T. R. (2015). *On your mark: Challenging the conventions of grading and reporting*. Bloomington, IN: Solution Tree.
- Guskey, T. R., & Bailey, J. (2001). *Developing grading and reporting systems for student learning*. Thousand Oaks, CA: Corwin.
- Guskey, T. R., & Jung, L. A. (2009). *Grading and reporting in a standards-based environment: Implications for students with special needs*. Paper presented at the Annual Meeting of the American Educational Research Association, San Diego, CA..
- Guskey, T. R., Swan, G. M., & Jung, L. A. (2011). Grades that mean something: Kentucky develops standards-based report cards. *Phi Delta Kappan*, 93(2), 52-57.
- Hattie, J. (1999). *Influences on student learning*. Inaugural professorial address, University of Auckland, New Zealand.

- Hattie, J. (2009). *Visible Learning: A synthesis of over 800 Meta-analyses relating to achievement*. New York: Routledge.
- Hattie, J. (2012). *Visible Learning for teachers: Maximizing impact on learning*. New York: Routledge.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112.
- Hooper, J., & Cowell, R. (2014). Standards-based grading: History adjusted true score. *Educational Assessment*, 19, 58-76.
- Kirschenbaum, H., Simon, S. B., & Napier, R. W. (1971). *Wad-ja-get? The grading game in American education*. New York: Hart.
- Kohn, A. (1994). Grading: The issue is not how but why. *Educational Leadership*, 52(2), 38-41.
- Kohn, A. (2011). The case against grades. *Educational Leadership*, 69(3), 28-33.
- Marzano, R. J. (2000). *Transforming classroom grading*. Alexandria VA: ASCD.
- Marzano, R. J. (2006). *Classroom assessment and grading that work*. Alexandria VA: Association of Supervision and Curriculum Development.
- Marzano, R. J., & Heflebower, T. (2011). Grades that show what students know. *Educational Leadership*, 69(3), 35-39.
- McElligott, J., & Brookhart, S. (2009). Legal issues of grading in the era of high-stakes accountability. In T. R. Guskey (Ed.), *Practical solutions for serious problems in standards-based grading*, (pp. 9-26). Thousand Oaks, CA: Corwin.

- McMillan, J. H. (2009). Synthesis of issues and implications for practice. In T. R. Guskey (Ed.), *Practical solutions for serious problems in standards-based grading* (pp. 105-120). Thousand Oaks, CA: Corwin.
- Muñoz, M. A., & Guskey, T. R. (2015). Standards-based grading and reporting will improve education. *Kappan*, 96(7), 64-68.
- National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform*. Washington, DC: U.S. Department of Education.
- North Carolina Department of Public Instruction. (n.d.). Improving the academic achievement of the disadvantaged. Retrieved from <http://www.ncpublicschools.org/program-monitoring/titleIA/>
- O'Connor, K. (2011). *A repair kit for grading: 15 fixes for broken grades*. Boston: Pearson.
- Ohlsen, M.T. (2007). Classroom assessment practices of secondary school members of NCTM. *American Secondary Education*, 36(1), 4-14.
- Olson, L. (1995). Cards on the table. *Education Week*, 23-28.
- Ravitch, D. (2010). *The death and life of the great American school system: How testing and choice are undermining education*. New York, NY: Basic Books.
- Reeves, D. B. (2004). The case against the zero. *Phi Delta Kappan*, 86, 324-326.
- Reeves, D. B. (2008). Effective grading practices. *Educational Leadership*, 65(5), 85-87.
- Reeves, D. B. (2011). Taking the grading conversation public. *Educational Leadership*, 69(3), 76-79.

- Reeves, D. B. (2012). The ketchup solution. *American School Board Journal*, 199(7), 35-36.
- Reeves, D. B. (2016). *Fast grading: A guide to implementing best practices*. Bloomington, IN: Solution Tree Press.
- Reeves, D. B, Jung, L., & O'Connor, K. (2017). What's worth fighting against in grading? *Educational Leadership*, 74(8), 42-45.
- Schimmer, T. (2016). *Grading from the inside out*. Bloomington, IN: Solution Tree Press.
- Scriffiny, P. L (2008). Seven reasons for standards-based grading. *Educational Leadership*, 62(2), 70-74.
- Smallwood, M. L (1935). *An historical study of examinations and grading systems in early American universities*. Cambridge, MA: Harvard University Press.
- Southern Regional Education Board. (n.d.). *Ten strategies for creating a classroom culture of high expectations* [PDF file]. Retrieved from https://www.sreb.org/sites/main/files/file-attachments/04v03_ten_strategies_0.pdf
- Starch, D., & Elliott, E. C. (1912). Reliability of the grading of high school work in English. *School Review*, 20, 442-457.
- Starch, D., & Elliott, E. C. (1913). Reliability of the grading of high school work in mathematics. *School Review*, 21, 254-259.
- Stiggins, R. (2000). Forward. In E. Trumbull & B. Farr (Eds.), *Grading and reporting student progress in an age of standards* (pp. ix-xi). Norwood, MA: Christopher-Gordon Publishers.
- Stiggins, R. J. (2001). The unfulfilled promise of classroom assessment. *Educational Measurement: Issues and Practice*, 20(3), 5-15.

- Stiggins, R., & Chappuis, J. (2005). Using student-involved classroom assessment to close achievement gaps. *Theory into Practice*, 44(1), 11-18.
- Summers, P. M. (2009). *Middle school grading practices and the ability to predict achievement on the Arkansas benchmark test* (Doctoral dissertation). Available from ProQuest Dissertations & Theses Global: (3372342)
- Thiele, J. (2016). *Grading practices and mindset development: The growth of both* (Order No. 10161494). Available from ProQuest Dissertations & Theses Global: The Humanities and Social Sciences Collection. (1842422357)
- Tomlinson, C. A., & McTighe, J. (2006). *Integrating differentiated instruction and understanding by design*. Alexandria, VA: Association for Curriculum and Supervision Development.
- Troxell, K. F. (2009). *The efficacy of teacher grades in predicting standardized testing outcomes* (Doctoral dissertation). Retrieved from ProQuest Dissertation and Thesis. (UMI No. 3353825)
- Trumbull, E., & Farr, B. (2000). *Grading and reporting student progress in an age of standards*. Norwood: Christopher-Gordon Publishers, Inc.
- United States Department of Education. (2018). *Improving basic programs operated by local educational agencies (Title I, Part A)*. Retrieved March 14, 2019, from <https://www2.ed.gov/programs/titleiparta/index.html>
- Vatterott, C. (2015). *Rethinking grading: meaningful assessment for standards-based learning*. Alexandria, VA: ASCD.
- Waltman, K. K., & Frisbie, D. A. (1994). Parent's understanding of their children's report card grades. *Applied Measurement in Education*, 7(3), 223-240.

- Webber, J., & Wilson, M. (2012). Do grades tell parents what they want and need to know? *Kappan*, 94(1), 30-35.
- 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.
- Wiggins, G. (1996). Honesty and fairness: Toward better grading and reporting. In T. Guskey. (Ed.), *Communicating student learning: The 1996 ASCD yearbook* (pp. 141-177). Alexandria, VA: Association of Supervision and Curriculum Development.
- Wiliam, D. (2011). *Embedded formative assessment*. Bloomington, IN: Solution Tree.
- Wiliam, D. (2016). The secret of effective feedback. *Educational Leadership*, 73(7), 10-15.
- Winger, T. (2005). Grading to communicate. *Association for Supervision and Curriculum Development*, 63(3), 61-65.
- Wormeli, R. (2006a). Accountability: Teaching through assessment and feedback, not grading. *American Secondary Education*, 34(3), 14-27.
- Wormeli, R. (2006b). *Fair isn't always equal: Assessment and grading in the differentiated classroom*. Portland, ME: Stenhouse.
- Wormeli, R. (2011). Redos and retakes done right. *Educational Leadership*, 69(3), 22-26.

Appendix A

Teacher Consent to Participate

Teacher Consent to Participate Form

Dear Participant,

The following outlines the purpose of the study and other relevant information that will help you decide if you would like to participate in the study. Please know that even if you give your consent to participate in the study, you still have the right to voluntarily withdraw from the study at any time without prejudice.

The central purpose of this qualitative study is to describe teachers' current classroom grading practices and explain the process by which teachers' grading practices develop and evolve throughout their teaching career. In addition, the study seeks to explore the various influences on teachers' grading practices. It is hoped that through this study a theory will be developed to, not only, explain the process by which teachers' grading practices develop and evolve throughout their teaching career, but also to equip stakeholders with more effective ways of influencing grading practices positively.

Data will be collected mainly through interviews, documents and surveys. All interviews will be recorded and later transcribed and coded. Pseudonyms will be used to protect the identity of all participants.

There are no known risks and/or discomforts associated with this study. However, one possible benefit of participating in the study is that participants gain an opportunity to contribute to the development of a theory on how teachers' grading practices evolve throughout their teaching career.

Your signature below will indicate that you have been provided with the information about the purpose of the study and that with this knowledge, you consent to participating in this study. Please do not hesitate to ask questions pertaining to the study before, during

or after you consent to participate. A copy of this consent form will be given to you to keep.

Date:

Name of the participant:

Signature of participant:

Name of the researcher:

Signature of researcher:

Appendix B
Teacher Interview Protocol

Teacher Interview Protocol

Time of Interview:

Date:

Place:

Interviewer:

Interviewee:

Position of Interviewee:

Project Description:

The central purpose of this qualitative study is to describe teachers' current classroom grading practices and explain the process by which teachers' grading practices develop and evolve throughout their teaching career. In addition, the study seeks to explore the various influences on teachers' grading practices. In this study *grading practices* refer to all the activities a teacher engages in that is related to and that leads to the assignment of grades/marks throughout a grading period.

Questions:

1. Think back to when you first started your teaching career. Tell me about your first experiences with grading and reporting on student learning.
 - a. *How did students and parents respond to your grading methods?*
 - b. *Describe your level of confidence that the grades you assign accurately reflect published academic standards. How has this changed over time?*
 - c. *In terms of grading and reporting on student learning, how have you grown since the start of your career?*
2. With your permission we will continue the interview by referring to grades you assigned to students throughout the first/second grading period of this school year.

When you are ready, please take a moment to retrieve grades from any previous

grading period for any one of your classes.

- a. Examine the final grades for students in any one of your classes. Based on your professional judgement and knowledge of your students, are there any grades that were assigned that now seem to misrepresent that students level of learning? Explain.
- b. You will describe the events/activities that led up to the final grade for any one student.

The following prompts will be used by the researcher if needed:

- i. To help guide your description, you may select a student from any one of your classes and describe/explain the various grades assigned and the events leading up to the assignment of each of those grades.
 - ii. *You can begin with a description of the activities you engage in before you start to teach a lesson (the planning process) followed by a description of the related lesson you taught. Your description can include details about the lesson and all the decisions and considerations you made leading up to the grades you eventually assigned. Your description could also address how you decide what is to be graded and what takes place after a grade is assigned.*
3. In the previous question, you described in detail all the practices you use when grading and reporting on student achievement.
 - a. Tell me about how you came to acquire those grading skills?
 - b. In what ways have your current grading practices changed from the start of your teaching career leading up to today?

4. Who or what has been most influential in the way you grade today?
5. How are the grades student receive related to the standards being taught?
6. Is there anything you would like to ask me before we end the interview?

I would like to remind you that all your responses will be held in strict confidence.

Thank you for your cooperation and participation in this study.

Demographic Data:

1. How many years of teaching experience have you had (including this current year)?
2. What is your current teaching assignment?
3. What level (e.g. BS, MS) of professional education have you attained?

Post Interview:

Upon conclusion and successful defense of my study, would you like to receive a summary of the findings?

Appendix C

Principal Interview/Questionnaire

Principal Interview/Questionnaire

Principal: _____

School: _____

Please circle the appropriate response after each statement:

1. Does your school/district have printed policies or procedures related to teacher grading practices? YES NO
2. If you answered “YES” to question # 1, please include a copy with the completed questionnaire or at the interview.
3. How frequently do you and the School Improvement Team (SIT) review grading practices? Describe any recent changes that have been made to the school’s grading policies, and the circumstances that led to those changes.
4. Education reform efforts over the past few decades have focused heavily on standards. Today, standards are well established for most subjects and state assessments used to gauge student mastery are also aligned to these standards.
 - a. How important is it that the grades assigned to students throughout the year provide information about student learning related to standards?
 - b. To what extent do the grades student receive support their learning on these standards?
5. The following statements represent expectations expressed in the _____ Grading practices/Guidelines which is posted on the district’s website. Please explain your understanding of each statement below.
 - a. *Academic grades shall be a content-based measure of what students are able to demonstrate.*

- b.** *Grades shall reflect performance that is consistent with demonstrated achievement on the goals of the North Carolina Standard Course of Study.*

6. To what extent are the grades assigned by teachers at your school consistent with these expectations?

7. Please indicate/comment on which of the following grading practices are being utilized as part of your school's grading policies or practices:

- A. Effort, participation, and compliance with rules are not included in grades.
- B. Academic achievement is not penalized due to student misbehavior.
- C. Grade reductions are applied when student work is submitted late.
- D. Extra credit points are included in grades.
- E. Multiple opportunities for assessment (retakes) are expected.
- F. Grade reductions are applied as a consequence for academic dishonesty.
- G. Student attendance concerns are reported separately from academic achievement and do not play a role in determining grades.
- H. Grades are organized and reported by standards or learning goals.
- I. Grades are assigned based on each student's performance compared to standards.
- J. Zeros are included in grades when student work is missing.
- K. The mean/average is used to determine a student's overall grade.
- L. Professional judgement and other measures of central tendencies (median/mode...) may be used in the determination of a student's overall grade.

M. Homework, practice and formative assessments are not used to determine grades, only summative evidence is used.

8. Based on the grades you've observed and number of students earning honor roll...
do you think that the grades teachers assign relate closely to students
learning/mastery of standards taught?
9. What changes (if any) might be implemented in your school/district grading
policies which would help to provide a better measure of assessing student
achievement?

Appendix D

District Leader Interview Protocol

District Leader Interview Protocol

Time of Interview:

Date:

Place:

Interviewer:

Interviewee:

Position of Interviewee:

Project Description:

The central purpose of this grounded theory study is to describe teachers' current classroom grading practices and explain the process by which teachers' grading practices develop and evolve throughout their teaching career. In addition, the study seeks to explore the various influences on teachers' grading practices. In this study *grading practices* refer to all the activities a teacher engages in that is related to and that leads to the assignment of grades/marks throughout a grading period.

Interview Questions:

Please review the document which represents the grading practices established by your school district. The document which is posted on the school district's website was created by a committee of elementary, middle and senior high teachers, principals, parents, students and central level administrators who met during the 2010-2011 school year.

1. The following statements represent expectations expressed in the _____ Grading practices/Guidelines. Please explain your understanding of each statement below.

- a. *Academic grades shall be a content-based measure of what students are able to demonstrate.*

b. *Grades shall reflect performance that is consistent with demonstrated achievement on the goals of the North Carolina Standard Course of Study.*

2. How are these expectations and policies monitored and how important do you think it is that these expectations be followed?
3. The _____grading practices/guidelines has a specified policy titled *Reteach/Retest* which only applies to high schools. What do you think is the reason for establishing this policy for high schools only? Do you think this should be extended to middle and elementary schools too?
4. How frequently does the district examine and make changes to the grading policies/guidelines?
5. What is the district's policy on the use of zeros?
6. What is the district's policy on the use of measures of central tendencies other than the mean?
7. Education reform efforts over the past few decades have focused heavily on standards. Today, standards well established for most subjects and state assessments used to gauge student mastery are also aligned to these standards.
 - a. How important is it that the grades assigned to students throughout the year provide information about student learning related to standards?
 - b. To what extent do the grades student receive support their learning on these standards?
8. What changes (if any) might be implemented in the current district grading policies which would help to provide a better measure of assessing student achievement?

Post Interview:

Upon conclusion and successful defense of my study, would you like to receive a summary of the findings?

☐ Yes

☐ No

Appendix E

Permission to Conduct Study

February 25, 2018,

Dear _____ :

I am currently pursuing a doctoral degree in Curriculum and Instruction at the Gardner-Webb University, Charlotte, NC.

I am interested in conducting a qualitative research study to describe teachers' grading practices and explain the process by which teachers' grading practices develop and evolve throughout their teaching career. I am therefore seeking permission to conduct interviews and gather various forms of archival data that would assist me in completing this study. All data published or reported will be kept anonymous.

At all stages of this research I will adhere to the guidelines of the university research ethics committee including anonymity of all reported data. I look forward to hearing from you and hope that it will be possible to use data from Title I Middle schools for this purpose. Any additional information that you may require can be provided upon request. If you have any questions concerning this research study please call/email Ingrid Rockhead at XXXXXX, email: XXXXXX_or Dr. Philip Rapp (Dissertation Committee Chair) at XXXX.

Please sign below to grant permission to carry out my research with teachers, principals, and district leaders at _____ Schools. Please keep one copy for your file and return the signed copy.

Signature

Superintendent, _____ Schools

Date

Sincerely,

Ingrid G Rockhead
Doctoral Candidate
Gardner-Webb University

Appendix F
Participant Profile

Participant Profile

Pseudonym	Grade Level	Subject	Experience/years
Hanna	8	Science	5-10
Parker	8	Math	10-15
Tasha	6	Math	5-10
Yazmine	7	Science	0-2
Linda	7	English	0-5
Sally	6	Math	5-10
Larry	8	Social Studies	0-2
Anders	8	English	0-2
Bailey	8	Math	0-5
Gad	7	Social Studies	15-20 years
John	7	Social Studies	0-2 years
Anna H	Middle grades (6-8)	Principal	>25 years
Dr. Bentley	N/A	District Leader	>25 years
Ken O'Connor	N/A	Author	N/A

Hanna is an 8th grade teacher of science and math with 6 years of teaching experience.

Prior to starting her teaching career, Hanna worked in the private sector. She later entered the teaching profession through the lateral entry pathway. She has spent all of her teaching career in the district.

Parker is an 8th grade teacher of mathematics with over 13 years of teaching experience.

She has spent all of her teaching career in the district.

Tasha is 6th grade teacher of Mathematics with over 6 years of teaching experience. She has spent all of her teaching career in the district.

Yazmine is a 7th grade lateral entry teacher of Social Studies and Science teacher. This is her first-year teaching.

Linda is a 7th grade teacher of English with over 5 years teaching experience. Prior to entering the teaching profession, she worked in the business world before joining the profession through the lateral entry pathway.

Sally is a 6th grade teacher of Mathematics with over 6 years of teaching experience. This is her second-year teaching in the district. Sally joined the teaching profession via the lateral entry pathway. Sally has taught 4 years in the elementary setting and this is her second-year teaching in a middle school.

Larry is an 8th grade teacher of Social Studies. This is his first-year teaching. He is a recent university graduate with a bachelor's degree.

Anders is an 8th grade teacher of English Language arts. He is a first-year teacher who recently graduated from college with a bachelor's degree.

Bailey is an 8th grade teacher of Mathematics with 5 years of teaching experience. She has taught Math for 4 years with the county/district and spent 1-year teaching in a neighboring school district. She has attained a master's degree in science.

Gad is an 8th grade teacher of social studies with over 18 years of teaching experience. Gad has taught for 8 years with the district. Gad is the holder of a master's degree.

John is a 7th grade teacher of social studies with 2 years of teaching experience. This is his second year working with the district.

Anna Holly is the Principal of one of the three middle schools at which most of the participants teach.

Dr. Ray is a district leader who is charged with the responsibility of overseeing academics for the district.

Ken O'Connor is an expert on grading and reporting, as well as the author of several books and articles related to grading and effective communication of student achievement. As a leading expert on the subject of grading, Mr. O'Connor provided data and insight to help resolve the third research question: What changes (if any) might be implemented in current grading practices which would provide a better measure of assessing student achievement?

Appendix G

Author/Grading Expert Interview

Author/Grading Expert Interview Protocol

Time of Interview:

Date:

Place:

Interviewer:

Interviewee:

Position of Interviewee:

Project Description:

The following outlines the purpose of the study and other relevant information that will help you decide if you would like to participate in the study. Please know that even if you give your consent to participate in the study, you still have the right to voluntarily withdraw from the study at any time without prejudice.

The central purpose of this qualitative study is to describe teachers' current classroom grading practices and explain the process by which teachers' grading practices develop and evolve throughout their teaching career. In addition, the study seeks to explore the various influences on teachers' grading practices. In this study *grading practices* refer to all the activities a teacher engages in that is related to and that leads to the assignment of grades/marks throughout a grading period.

Please indicate below your preferred level of privacy for this interview.

- ☐ I would like to remain anonymous.
- ☐ As an author/expert in the field of grading, I consent to having my name disclosed for the purpose of this research.
- ☐ Other: (*you may specify below any other way you would like the data you provide to be treated*) _____

Your signature below will indicate that you have been provided with the information about the purpose of the study and that with this knowledge, you consent to participating

in this study. Please do not hesitate to ask questions pertaining to the study before, during or after you consent to participate. A copy of this consent form will be given to you to keep.

Date:

Signature of participant:

Signature of researcher:

Questions:

Please review the document which represents the grading practices established by a small school district in North Carolina. The document which is posted on the school district's website was created by a committee of elementary, middle and senior high teachers, principals, parents, students and central level administrators who met during the 2010-2011 school year.

1. What aspects of the grading guidelines do you think serve as a positive influence teachers' grading practices and should be retained? What would be your reason for preserving those parts of the grading policy?
2. If you could change any aspect of the districts grading policies, what would you change? What would be your reason for changing it?

Post Interview:

Upon conclusion and successful defense of my study, would you like to receive a summary of the findings?

☐ Yes

☐ No

Appendix H

District Grading Policy

Schools Grading Guidelines

Grades 3-12

General Guidelines

Purpose: To provide direction to all teachers in expected grading practices

Expectations:

- Grading practices shall be reviewed annually by the principal and School Improvement Team.
- Academic grades shall be a content-based measure of what students are able to demonstrate.
- Grading shall provide appropriate and accurate feedback to the students and parents.
- Grades shall reflect performance that is consistent with demonstrated achievement on the goals of the North Carolina Standard Course of Study.

Guiding Documentation

Purpose: Written documentation of how a grade is derived is essential for communicating with parents.

Expectations:

Elementary (3-5)	Middle (6-8)	High (9-12)
<ul style="list-style-type: none"> • Each grade level within a school shall be responsible for providing a written document to parents explaining their grading guidelines. • Each grade level's document must be approved by the principal. • Upon approval, the grade level document shall be sent home and signed by parents within the first two weeks of school. • Teachers shall place the signed document in the Student Accountability Folder each year and purge at the end of the year. • This document shall be given to all students at their time of enrollment. 	<ul style="list-style-type: none"> • All teachers are required to have a syllabus for each course taught. • The syllabus must be approved annually by the administration. • The syllabus must be communicated annually to students and parents via website and/or handout. • The syllabus must include the breakdown of grading plan to include percentages or points. • The syllabus must include a list of classroom procedures. 	<ul style="list-style-type: none"> • All teachers shall be required to give the Principal or his designee and students a course syllabus that contains at least the course title, course description, general curriculum topics to be covered, materials, grading practices which may include an example of a calculated student grade, classroom procedures including behavior management, teacher contact information. Course syllabi must be distributed to all students within the first week of the semester, or student enrollment. Principals shall review each course syllabus annually.

Grade Composition

Purpose: Grades shall reflect student performance in various subject areas aligned with the North Carolina Standard Course of Study. The make-up of the student grade shall be based on various assignments with differing percentages. The assignments shall carry different weights based on the importance of the activity, time spent on the activity and the degree of involvement.

Expectations:

Elementary	Middle	High
<ul style="list-style-type: none"> • Reading and writing grades – The weights of each shall be reading 70 percent and writing/word study 30 percent. • In reading and math, at least three (3) must be test grades that are reflective of their proficiency on the North Carolina Standard Course of Study. • Homework shall not be included in grades. • If a child is working below grade level, it shall be noted in the comments section of the report card. • Academic grades shall not be reduced for misconduct or for punishment. • Grades shall be based on percentages, not points. • The final Language Arts grade shall be based upon a minimum of ten (10) reading assignments and 	<ul style="list-style-type: none"> • The nine weeks grade shall be determined by the following: <ol style="list-style-type: none"> 1. 40 percent - tests (major - projects, papers and performances) 2. 30 percent independent work - (quizzes, classwork, common assessments, etc.), 3. 20 percent guided work - (group activities, teacher led activities, etc.) and 4. 10 percent homework • A minimum of three (3) tests shall be given during each grading period for all middle school courses. In addition to test grades, a minimum of five (5) guided assignments and four (4) independent assignments shall be given in each grading period for all middle school courses. • Academic grades shall not be reduced for misconduct or for punishment. • Teachers shall supply a copy of their grading procedures to the principal, students and parents. • Teachers shall work with Professional Learning Communities to determine interventions and monitor progress to ensure mastery of content knowledge. • Teachers shall enter grades in PowerTeacher weekly. 	<ul style="list-style-type: none"> • The six weeks grade shall be determined by counting tests (major projects, major papers, etc.) 60 percent, quizzes 20 percent and, daily/class grades and/or homework 20 percent. • All students shall take a final exam in all courses. The exam grade shall count 20% of the final course grade. The final exam is a cumulative exam created by the State, District, or teacher and shall include all standards in the course curriculum. • Each six weeks grade shall be made up of a minimum of two (2) major test grades (may include major projects, essays, research papers, performances) which shall count 60 percent of the grade. • Each six weeks grade shall be made up of a minimum of four (4) quizzes (only covers part of a unit of study or a specific step in a project or performance) which shall count 20 percent of the grade. • Each six weeks grade shall be made up of a

<p>four (4) writing assignments.</p> <ul style="list-style-type: none"> • The final math grade shall be based upon a minimum of ten (10) assignments. • The science grade shall be based upon a minimum of four (4) assignments. • The social studies grade shall be based upon a minimum of four (4) assignments. • Teachers shall enter grades in PowerTeacher weekly. • School level professional learning communities shall help determine an appropriate number of additional assignments to be given for each discipline. • The teacher shall adhere to the IEP, the 504 and any Tier Plan. 	<ul style="list-style-type: none"> • School level professional learning communities shall help determine an appropriate number of additional assignments to be given for each discipline. • The teacher shall adhere to the IEP, the 504 and any Tier Plan. 	<p>minimum of six (6) daily/class and/or homework grades (may include warm-ups or bell work) which shall count 20 percent of the grade. Exceptions to this shall be made for all non-core classes except foreign language.</p> <ul style="list-style-type: none"> • AP/Honors courses shall not be graded differently. Content and expectations shall be more rigorous. • Fine Arts, PE and ROTC shall include grades for performance. • Academic grades shall not be reduced for misconduct or for punishment. In accordance with BOE Policy #4310, students will be subject to consequences outlined in Rule 25 of the Student Code of Conduct for cheating, plagiarism, etc. • Teachers shall enter grades in PowerTeacher weekly. • School level Professional Learning Communities shall help determine an appropriate number of additional assignments to be given for each discipline. • The teacher shall adhere to the IEP, the 504 and any Tier Plan.
---	---	--

Homework Assignments

Purpose: Homework is important in supplementing classroom instruction and furthering the goals of the educational program. It reinforces learning and stimulates independence, responsibility and self-direction.

Expectations:

Elementary	Middle	High
<ul style="list-style-type: none"> • Homework assignments shall be designed to provide 	<ul style="list-style-type: none"> • Homework assignments shall be designed to provide 	<ul style="list-style-type: none"> • Teachers shall work with Professional Learning Communities to determine amount and frequency of homework and it shall

<p>practice, preparation, or extension to curriculum taught in the North Carolina Standard Course of Study.</p> <ul style="list-style-type: none"> Homework shall not be longer than the grade level X 10 minute rule (ex. A second grader would have no more than 2 times 10 minutes); excluding reading practice. Total time includes completing unfinished class work. Homework shall not be counted as a grade. Teachers shall avoid using homework as punishment. 	<p>practice, preparation, or extension to curriculum taught in the North Carolina Standard Course of Study.</p> <ul style="list-style-type: none"> Teachers shall work with Professional Learning Communities to determine amount and frequency. Teachers shall assign a reasonable amount of homework and it shall count no more than 10 percent of the final grade. Students shall receive feedback on homework. Teachers shall avoid using homework as punishment. 	<p>count no more than 20 percent of the final grade. Also included in this 20 percent are class and daily grades.</p> <ul style="list-style-type: none"> Homework shall be meaningful and relevant to the North Carolina Standard Course of Study. Teachers shall avoid using homework as punishment.
--	---	---

Interim/Report Card Grades

Purpose: Every student shall be given the opportunity to earn a passing final grade.

Expectations:

Elementary	Middle	High
<ul style="list-style-type: none"> Teachers shall not give a final grade below 50 on the report card for all marking periods. The teacher shall utilize the appropriate PowerTeacher comment on the report card. Actual grades shall be reflected on assignments that are returned to students and in the grade book. 	<ul style="list-style-type: none"> Teachers shall not give a grade below 50 on the report card for all marking periods. The teacher shall utilize the appropriate PowerTeacher comment on the report card. Interim reports shall reflect actual grades. The teacher shall maintain frequent communication with student and parent. 	<ul style="list-style-type: none"> *Credit Recovery and *Repeating a Course for credit will be available to students who fail a course. Teachers shall not give a grade below 50 on the report card for the first and fourth six weeks grading periods. Interim reports shall reflect actual grades. The teacher shall maintain frequent communication with student and parent. *Grade Recovery (as described below) will be utilized for students who have not successfully completed

<ul style="list-style-type: none"> Interim reports shall reflect actual grades. The teacher shall maintain frequent communication with student and parent. 		assignments/homework for a particular course while in progress.
--	--	---

***Definitions:**

Students enrolled in Grade Recovery have the opportunity to receive a passing grade in a course they did not pass on the most recent report card. Students complete parts of the course curriculum and/or assignments that contributed to them not passing. When a student successfully completes the Grade Recovery their original report card grade will be changed to a passing mark of 60.

Credit Recovery is defined as a block of instruction that is less than the entirety of the North Carolina Standard Course of Study for that course. When Credit Recovery is exercised, the original record of the course being completed and failed will remain on the transcript. A grade of pass or fail will be given for each credit recovery course. The mark will not affect the student's GPA. Credit Recovery enables students to recover course credit; it is not intended for grade replacement.

Repeating a Course will be used to refer to a high school course repeated via any delivery method when the entire North Carolina Standard Course of Study for the course is being taught to the student for a second time. Students participating in Repeating a Course for credit will receive a grade and take the associated end-of-course assessment, if appropriate.

Core/Encore Classes

Purpose: All middle school courses are of equal importance in promoting a balanced curriculum for students.

Expectations:

Elementary	Middle	High
N/A	<ul style="list-style-type: none"> All middle school courses shall carry the same graded weight regardless of the subject or academic level. All middle school courses shall establish procedures for grading and assessing their students. This includes but is not limited to Health and PE, Band, Chorus, Art, Foreign Language and CTE courses. The grades shall reflect what a student knows and is able to do. All classes shall use a variety of assessments to measure learning. Tests shall include but are not limited to performances, physical activity assessments, content assessments and performance based projects. 	N/A

Reteach/Retest

Purpose: High School students shall be given the opportunity to participate in reteach/retest sessions.

Expectations:

Elementary	Middle	High
------------	--------	------

N/A	N/A	<ul style="list-style-type: none"> • The reteach/retest program in _____ high schools is a strategy to be utilized in working with student mastery of coursework. This program is intended to be a complement to the regular instructional day and shall be used to increase student mastery of subject matter. • A reteach session is not mandatory before the retest is administered, but can be required by the teacher if he/she believes it to be necessary and beneficial to the student. • Benchmarks, formative assessments and final exams shall not be in the reteach/retest program. • Reteach/retest can be a part of a student's makeup of work missed due to an absence (this can be addressed within the school-specific guidelines). • Students shall be able to re-test any major assessment one (1) time. • Any student who makes below a 75 is eligible to re-test. • The highest grade a student may make on a re-test is 75. If the re-test grade is lower than the original test grade, the student shall receive the higher of the two grades. • Tests that may be retested are those which comprise the 60 percent of the six weeks grade (major assessments) with exceptions listed above. Entire projects may not be considered part of the retest process. Teachers may allow students to "retest" a portion of a project, therefore positively affecting the overall grade assigned. • The reteach-retest guidelines apply to all courses offered in __, including honors, Advanced Placement (AP) courses, and Gaston Online courses. However, the ____ Grading Guidelines do not apply to CCP and NCVPS courses. • It is recommended that students complete the reteach/retest process for a major assessment within 10 days of original assessment administration. • Students who qualify for the exceptional children's program may earn higher than a 75 if reteach/retest is included as a modification in the IEP. • Schools must publish the district guidelines and school specific guidelines regarding days and times for retesting to parents, students, and staff in a written document (preferably the student handbook).
-----	-----	---

Revised August 2015