



SCHOOL of
GRADUATE STUDIES
EAST TENNESSEE STATE UNIVERSITY

East Tennessee State University
Digital Commons @ East Tennessee
State University

Electronic Theses and Dissertations


Student Works

5-2023

Selecting Classroom Grading Practices: A Qualitative Study Exploring Teacher Experiences

Tallye Gass

Follow this and additional works at: <https://dc.etsu.edu/etd>

 Part of the [Educational Leadership Commons](#), [Educational Methods Commons](#), and the [Elementary and Middle and Secondary Education Administration Commons](#)

Recommended Citation

Gass, Tallye, "Selecting Classroom Grading Practices: A Qualitative Study Exploring Teacher Experiences" (2023). *Electronic Theses and Dissertations*. Paper 4217. <https://dc.etsu.edu/etd/4217>

This Dissertation - unrestricted is brought to you for free and open access by the Student Works at Digital Commons @ East Tennessee State University. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of Digital Commons @ East Tennessee State University. For more information, please contact digilib@etsu.edu.

Selecting Classroom Grading Practices: A Qualitative Study Exploring Teacher Experiences

A dissertation

presented to

the faculty of the Department of Education Leadership and Policy Analysis

East Tennessee State University

In partial fulfillment

of the requirements for the degree

Doctor of Education in Educational Leadership

by

Tallye Wehenkel Gass

May 2023

Dr. Pamela Scott, Chair

Dr. Stephanie Barham

Dr. William Flora

Dr. Virginia Foley

Keywords: grades, grading practices, grading systems

ABSTRACT

Selecting Classroom Grading Practices: A Qualitative Study Exploring Teacher Experiences

by

Tallye Gass

The purpose of this qualitative research was to understand the experience of selecting grading practices for a high school math or English classroom. There is no existing research related to understanding this phenomenon. This phenomenological study included three participants who were purposefully selected based on specific criteria: They must teach math or English in grades 9, 10, 11, or 12 and have at least three years of teaching experience. The participants were interviewed using videoconferencing software, and the interviews were recorded and transcribed. Open and axial coding was used to analyze the data. The theoretical framework applied for the interpretation of the findings was Social Learning Theory and Social Cognitive Theory. The participants shared thick, rich data from which eight themes emerged: Communication among stakeholders, student effort, value of learning, minimal education in grades and grading practices, external pressures, grades as motivators, teacher expectations, and reliability and validity in grading practices. Through this research, it became understood that there are many facets of this experience that are all inter-related. Based on the research recommendations were made for professional practice and future research to ignite positive change around this phenomenon.

Copyright 2023 by Tallye Gass

All Rights Reserved

DEDICATION

Every word on these pages is dedicated to my family: I owe them with every breath of my body. This is for my Grandpa, who made me believe learning was magical and could fuel my dreams, and for my Nana, who called me “Shug” and was magnificent in her strength and resolve. It is for my Maudie, who gave me everything because she never had it, and my Hilbert, who believed I walked among the stars and showered me with the Heavens and all their light. It is for Aunt Wanda, who left me entirely too soon, but gave me all the love she had while she was here. It is for my wonderful in-laws who raised the amazing man I was lucky enough to marry and have supported me endlessly. It is most certainly for my parents, who put up with way too much from me but still love me more than all the grains of sand swirling in the oceans and support my crazy ideas with no questions. Thank you for being marvelous and cool parents who gave me all the books my heart desired and gave me room to grow even when it hurt.

To my children Grace, Jacob, and Cate: thank you for allowing me to do this. It has taken time from each of you over this journey, but I hope I have made you proud. When I say you are my reason for waking every day, it is my whole truth. You are each kind, talented, and loving. You are each wise, funny, and all the best things in life. I can’t wait to see the goodness you splash across the world.

Lastly, to my husband. There are simply no words. You put me back together when I was shattered and walked by my side through it all. Everything I do is only possible because you are by my side. I am eternally grateful to spend my forever with you. You are the mate to my soul in this life and beyond.

ACKNOWLEDGEMENTS

I would like to thank my committee members Dr. William Flora, Dr. Virginia Foley, and Dr. Stephanie Barham for being experts in their fields and sharing this knowledge with others. It is through their courses that I began to see the beginnings of research evolve from the seed of curiosity into a full body of work. A special thanks for my chair, Dr. Pamela Scott, who has been by my side from my first inquiry email about the program. When I doubted myself, her confidence kept me moving forward. It will always be an honor to have completed this journey under her guidance.

I would also like to thank Dr. Suzanne Bryant, who has unknowingly served as my mentor for over a decade. She encouraged me to pursue this goal and assured me I could do it. She is a champion for all students, both young and old, and a true advocate for public education.

I must also thank my participants without whom none of this would have been possible. Their willingness to be open and share their experiences is the heart of this study.

A final thank-you to all the teachers who put up with me and pushed me over the years. I was not an easy student, but I appreciate your faith in me and thank you for not giving up on me. Mrs. Barger, thanks for telling me I was a good writer. I believed you. I pray I am that teacher for someone, too.

TABLE OF CONTENTS

ABSTRACT	2
DEDICATION	4
ACKNOWLEDGEMENTS	5
TABLE OF CONTENTS	6
LIST OF TABLES	12
Chapter 1. Introduction	13
Statement of the Problem	13
Purpose Statement	13
Research Question	14
Supporting Research Questions	14
Significance of the Study	14
Definition of Terms	15
Limitations and Delimitations	17
Summary	18
Chapter 2. Review of the Literature	19
Foundations of Education in the United States	19
Comprehensive Public High Schools in the United States	21
Grades	25
Origin of Grades	25
Early Grading	26

Purpose of Grades	27
Homework.....	28
Assessments	31
Formative Assessments.....	31
Summative Assessments.....	32
Standardized Assessments.	33
Measurement.....	33
Validity	34
Reliability.....	35
Equating of Scores	38
Square Root and Cube Root Calculations.....	40
Grading Systems	41
Criterion-Referenced Grading System.....	41
Norm-Referenced Grading System.....	41
Standards-Based Grading System.....	43
Percentage Grading System	43
Grading Scales	45
Grading Practices	46
Grading as Moral Responsibility	48
Mindset, Self-Concept, and Mental Health.....	49

Mindset.	49
Self-Concept.	51
Mental Health.	52
Academic and Non-Academic Factors in Grading	52
Grades as Motivators or De-motivators.....	55
Grades as Punishment or Reinforcement.....	56
Administrative and Parental Expectations	56
Trauma	57
The Brain.	58
Impact of Trauma on Learning.	59
Bias	61
Teacher Perception of Student Ability.....	61
Race	62
Gender.....	64
Familiarity of Student.	65
Local Grading Policy	65
Theoretical Framework.....	69
Summary	71
Chapter 3. Methodology	73
Research Questions.....	73

Supporting Research Questions	73
Qualitative Research Design.....	73
Site Selection	75
Participants.....	75
Data Collection Strategies.....	76
Interviews.....	77
Data Analysis Strategies	77
Assessment of Quality and Rigor	78
Trustworthiness.....	78
Thick Description.....	79
Triangulation.....	79
Reflexivity.....	79
Ethical Considerations	80
Summary	80
Chapter 4. Findings.....	81
Purpose of the Study	81
Research Questions.....	82
Data Collection and Data Analysis	82
Framework	82
Findings.....	83

Research Question One.....	85
Communication Among Stakeholders	85
Student Effort.....	86
Value of Learning	87
Research Question Two	88
How to Grade and Implement Grading Practices	89
Research Question Three	90
How to Grade and Implement Grading Practices	90
Research Question Four.....	91
External pressures	91
Additional Themes.....	94
Grades as Motivators	94
Teacher Expectations	96
Reliability and Validity in Grading Practices	97
Summary.....	98
Chapter 5. Conclusions	100
Discussion.....	100
Research Question One.....	100
Communication Among Stakeholders	101
Student Effort.....	101

Value of Learning	102
Theoretical Framework.....	103
Research Question Two	103
Theoretical Framework.....	104
Research Question Three	104
Theoretical Framework.....	105
Research Question Four.....	106
Theoretical Framework.....	107
Additional Themes.....	107
Grades as Motivators	107
Teacher Expectations.....	108
Reliability and Validity in Grading Practices	109
Theoretical Framework.....	109
Implications for Practice	110
Recommendations for Future Research.....	111
Summary	112
REFERENCES	113
APPENDIX: INTERVIEW QUESTIONS	137
VITA.....	139

LIST OF TABLES

Table 1. Percentage Score Example..... 44

Table 2. State of Tennessee Grading Scale for Grades 9-12 45

Table 3. District Grading Policy Analysis 67

Table 4. Summary of Emergent Theme Categories..... 84

Chapter 1. Introduction

Grades are a necessary and critical component of the educational process. It is essential they accurately convey the information they are intended to communicate. Grades can influence admission to post-secondary education institutions and access to scholarship monies (Adu-Mensah, 2018; Randall & Engelhard, 2010). Through the lens of validity, grades should focus on communicating the level of knowledge the student has demonstrated and not a reflection of behavior or effort.

Grades serve different but similar functions for students, parents, and teachers. Engelhard (2010) stated teachers use grades to compare the knowledge of current students, to determine the readiness of incoming students, and to communicate to students' and parents' understanding of the content. The students' responsibilities are to complete the tasks assigned, study the appropriate materials, and communicate their level of understanding on assessments (Tomlinson, 2005). Parents depend on students to meet their duties and teachers to assess and communicate student knowledge appropriately.

Statement of the Problem

The purpose of grades is to communicate feedback to students and parents on student understanding of the prescribed content. Existing research has demonstrated grades are a mixture of objective and subjective factors, leading to a lack of validity. Existing research is primarily quantitative; therefore, additional qualitative research is needed to understand the experiences educators have when selecting a grading system for their classroom.

Purpose Statement

The purpose of the phenomenological study was to understand the educator experience of choosing grading practices for math and English courses East Tennessee high schools.

Research Question

To understand the educator experience, the researcher developed a singular research question to guide the study: What are the lived experiences of high school math and English teachers when choosing grading practices in their classrooms?

Supporting Research Questions

Several research questions to answer the essential research question were developed.

RQ1: How do teachers perceive the use of grades in secondary math and English classrooms?

RQ2: What experiences are associated with choosing a classroom grading practice for secondary math and English classrooms?

RQ3: What experiences have teachers had with professional development centered on grading practices and policy for the classroom?

RQ4: What external influences affect teacher choice in classroom grading practices?

Significance of the Study

Grades have traditionally been an accepted education element influencing grade-level promotions and credits earned toward high school graduation. The National Center for Education Statistics (2022) documented the United States national adjusted cohort graduation rate at 86% for the 2018-19 school year. The highest adjusted cohort graduation rate was 92%, and the lowest adjusted cohort graduation rate was 69%. Grades are one of the most common reasons for students dropping out of high school and are powerful determinants to student success in and beyond school (Azzam, 2003; Dalton, et al., 2009).

Additionally, grades can determine access to post-secondary education for colleges and universities (Adu-Mensah, 2018; Randall & Engelhard, 2010). The high school grade point

average in Tennessee is one of two measures to determine eligibility for the Tennessee HOPE scholarship (Tennessee Department of Education, n.d.-b).

The body of research on grading systems has thus far focused on the reliability and validity of the grades. Reliability is defined as “the stability of the measured values obtained in repeated measurements under the same circumstances using the same measuring instrument” (Sürücü & Maslckçi, 2020, p.2). Where reliability focuses on the consistency of the measure, validity addresses the accuracy of the measure (Haladyna, 2019). Previous research has focused on the grading system, but a deeper investigation is needed to understand its selection.

Definition of Terms

1. *504 Plan* is defined as an individual plan that provides students with disabilities equitable access to learning (Jones, 2022).
2. *Advanced Placement course* is defined as high school courses that are taught at the rigor of a college course and for which students can receive college credit if they pass the Advanced Placement exam at the end of the course. This allows students to potentially take more advanced courses when they begin post-secondary schooling. These courses are run by the College Board (Edwards, 2021).
3. *Assessment* is defined as a method or toll used to “evaluate, measure, and document the academic readiness, learning progress, skill acquisition, or education needs of students” (The Glossary of Education Reform, 2015, para. 1).
4. *Bias* is defined as “a prejudice in favor of or against one thing, person, or group compared with another usually in a way that’s considered to be unfair” (University of California San Francisco Office of Diversity and Outreach, 2023, para. 1).

5. *Comprehensive high school* is defined as a secondary school where all students take “common” courses, such as English, math, science, and history, in addition to choosing from additional college preparatory, vocational, or self-interest courses (Copa & Pease, 1992; Wraga, 2000).
6. *Criterion-referenced grading systems* are defined as a grading system where a grade, usually a letter grade, is assigned based on a fixed, numerical scale (International Affairs Office, 2008).
7. *Economically Disadvantaged* in Tennessee is defined as a student who meets poverty eligibility guidelines to receive benefits from Supplemental Nutrition Assistance Program (SNAP), Temporary Assistance for Needy Families (TANF), and/or Head Start. They can also be homeless, a migrant, runaway, or in foster care (Tennessee Education Glossary, n.d.).
8. *End of course assessment* is defined as standardized exams administered at the end of specific high school courses that are designed to “assess true student understanding and not just basic memorization and test-taking skill” (Tennessee Department of Education, n.d.-c).
9. *Formative assessment* is defined as a low-stakes assessment worth little to no points that provides ongoing feedback to students and teachers about their current understanding of the learning (Carnegie Mellon University Eberly Center, 2023).
10. *Grades* are defined as a measurement used to quantify learning and intellectual progress using objective criteria (University of South Carolina, n.d.-b).
11. *Grading practices* are defined as a system used to determine a grade (University of South Carolina, n.d.-a).

12. *Graduation requirements* are defined as the required courses that must be completed with a passing grade to receive a high school diploma (Tennessee Department of Education, n.d.-a).
13. *Individualized Education Plan* is defined as a personalized plan or program developed for a student with an identified disability to support the student in meeting identified learning goals and provide educational benefits (Individuals with Disabilities Education Act, 2004; Jones, 2022).
14. *Norm-referenced grading systems* are defined as a grading system where a pre-established formula determines a grade with percentage ranges and students compete against one another. For example, the top 10% of the class would get an A, and the following top 20% would receive a B (International Affairs Office, 2008).
15. *Standards* are descriptions of the learning expectations for specific content areas in individual grade levels (Edglossary, 2014).
16. *Standards-based grading* is defined as a grading system where the teacher reports the student level of knowledge of a set of standards (Marzano, 2020).
17. *Summative assessment* is defined as an assessment that is usually high-stakes (high point value and/or weight) and is intended to evaluate the comprehensive learning at the end of the instructional period (Carnegie Mellon University Eberly Center, 2023).

Limitations and Delimitations

There were two limitations to this study. The sample size did achieve saturation; however, it was small. The second limitation was the personal bias of the researcher. The researcher had prior personal and professional experiences related to grading that could affect their ability to be impartial.

There were two delimitations to this study. Elementary teachers were not included in the study. These teachers may have had different experiences with selecting grading practices. The second delimitation was that test data was not included in the study. This would not provide any data related to the experience teachers had when choosing a grading practice.

Summary

The study is organized into five chapters. Chapter 1 includes the introduction, statement of the problem, purpose statement, research questions, the significance of the study, definition of terms, limitations and delimitations, and a summary of the study. Chapter 2 is a comprehensive review of the literature related to the history of secondary schools, the comprehensive high school model, grading practices, and the effects of grading practices and grades. Chapter 3 explains the research methodology. Chapter 4 details the data analysis and its findings. Chapter 5 summarizes the findings, includes implications of the research, and recommendations for future research.

Chapter 2. Review of the Literature

Grades and grading systems have been an integral part of the education system for over one hundred years (Guskey & Brookhart, 2019). While assessments used for grading have been extensively studied, qualitative research concerning teacher grading practices and their selection process is less extensive (Brookhart et al., 2016). Existing research primarily focused on the impact of grades and grading systems on students and the teaching and learning process (Duke & Canady, 1991; Guskey & Brookhart, 2019). Literature reviewed for this chapter focused on understanding the foundations of education in the United States; comprehensive high schools in the United States; grades; grading systems; the characteristics of criterion-referenced, norm-referenced, standards or mastery-based grading systems; bias in grading; and the effects of grades and grading systems on students and learning.

Foundations of Education in the United States

During the settlement of the colonies, education in the United States was a combination of the various organizations that immigrants had brought with them from their homeland; however, education was not a priority for many families during this time because they were focused on surviving and establishing themselves as new residents of the country (Copa & Pease, 1992; Gutek, 1991). During this time, the primary role of the school was to shape the moral fabric of the child (Vatterott, 2015). Teachers were expected to teach children to read so they could read the Bible, and the students were also expected to learn writing and arithmetic (Guskey, 2015; Vatterott, 2015).

The first attempt at compulsory, organized public schooling by a government was in 1642 when the colonial government passed the Massachusetts School Law of 1642 (Crofton, 2020). This law directed families to teach their children or apprentices to read and understand the laws

of society and faith at that time. The Massachusetts General School Law of 1647, commonly known as “The Old Deluder Satan Law,” required any town with 50 or more households to appoint and pay the wages for a teacher. Towns with more than 100 households had to establish a grammar school to prepare students for university (Copa & Pease, 1992; Crofton, 2020). These laws were not enforced with any regularity, and the penalties were not consequential, but they birthed the concept of publicly funded schools.

As America gained its independence during the Revolutionary period, leaders of this time, such as Benjamin Franklin, Thomas Jefferson, and Noah Webster, sought to design an educational system that rejected the European worldview (Guttek, 1991). This included a secular preference for education and a civic education to prepare students for future political responsibilities and shape America's cultural character. Thus began the foundations of national, systematic, and publicly funded education (Kober & Rentner, 2020).

Although the earliest secondary school had formed in 1635 with the founding of The Boston Latin Grammar School, it was the mid-1800s before high schools would be required and funded by taxes (Paterson, 2021). The first industrial revolution caused many immigrants to move to larger cities where this revolution was taking place, and leaders of these cities determined that children would need schooling (Guttek, 1991).

Common, or elementary schools, were publicly funded at this time; however, some citizens of Kalamazoo, Michigan believed anything beyond that should be paid for by the person or their family (Timmerman, 2012). In 1874, petitioners filed a lawsuit challenging the collection of taxes to support secondary school. The 1874 Kalamazoo School case in Michigan established that secondary school, or high school, should be publicly funded (Russo, 2008). This ruling increased the popularity of secondary schooling.

Comprehensive Public High Schools in the United States

As wages increased after the start of the 20th century, more working-class families sent their children to school (Library of Congress, n.d.). Before this, many children were responsible for helping support their families, and school was not an option for them. The high school structure as it is known today was essentially in place by the mid-1920s (Guttek, 1991). The most common model was four years of school covering grades 9 through 12, culminated with a diploma, and offered four primary programs: college preparatory, business, industrial vocation, or general education for those not pursuing a specific vocation or attending additional schooling.

Secondary school enrollment grew exponentially during this time. Between 1890 and 1930, the number of high school students increased by 1234.7%, from 359,949 to 4,804,255. By 1940, there would be 7,123,009 high school students between the ages of 14 and 17. Many communities, towns, and cities were building schools to provide access to secondary education for students. This propelled the conversation for a unified vision of education in the United States (Mirel, 2006).

Critics argued secondary school was too focused on university needs (Copa & Pease, 1992). This criticism was the foundation for the Commission on the Reorganization of Secondary Education (Kridel, 2010). The committee's report was published in 1918, and it argued for comprehensive secondary education with a focus on seven principles: "(a) health, (b) command of fundamental processes, (c) worthy home membership, (d) vocation, (e) citizenship, (f) worthy use of leisure, and (g) ethical character (Copa & Pease, 1992, p. L-10).

Comprehensive high schools in the early 1900s aimed to provide equal opportunities for education through courses in the liberal arts (Mirel, 2006). Unification was also a critical component of the argument for comprehensive secondary schools (Copa & Pease, 1992). The

purpose of the unification was to house students under the same roof while they were still specializing in different vocations, such as academics, clerical, agriculture, or household arts (Copa & Pease, 1992). Students would be unified by taking a series of common courses and housing students all in one building. Students took additional classes that were tied to specialization (constants), vocations (variables), and electives (Copa & Pease, 1992; Wraga, 2000). Specializations focused on subjects such as agriculture, business, or household arts. Vocations were specific training, and electives were courses of personal interests that were nonvocational. This combination of academic and vocational education housed in the same building was a hallmark of comprehensive education (Wraga, 2000).

The comprehensive education model had to be defended against critics who favored a less academic focus. Some educational leaders stated students did not have the “intellectual abilities” (Mirel, 2006, para. 14) for academically rigorous courses. They cited an increase in the number of students from immigrant families and the economic collapse of 1929 as the reasons for students being less able to succeed (Gutek, 1991; Mirel, 2006). While these critics did not successfully dismantle comprehensive high schools, the impact of the disparagement would continue to be felt for the coming decades in continued arguments for comprehensive programming to be separated by student interest and ability (Wraga, 2000).

Despite the ongoing call to dismantle comprehensive education, the comprehensive secondary schooling model remained unchanged in the second half of the 20th century. Course offerings expanded, and content was modified as standards were added, but the skeletal model remained intact (Wraga, 2000). As the responsibility of secondary school became increasingly custodial in the earlier part of the century, many parents depended on schools for non-academic resources (Mirel, 2006).

Critics argued comprehensive secondary schools offered academic content that was not rigorous enough; however, supporters of the model during the 1960s believed increasing academic demands would lead to decreased equality because it would create “a wave of dropouts” (Mirel, 2006, para. 29). At this time, schools were thought to have a democratic mission to provide greater education equality, and preventing students from dropping out would meet this goal (Mirel, 2006).

In 1959, James Conant, a retired Harvard University president, published his study on the American high school (Conant, 1959). He made three recommendations for improving high schools. First, there should be one counselor for every 250 to 300 students. Next, there should be an individual track for each student with no student being labeled as vocational or college preparatory. Finally, he recommended requiring general education courses for all students. These included four years of English, three or four years of social studies (two years of which must be history, and one year must be American history), and a year-long course in some form of American government during their senior year. Students would also take one year of math and one year of science. He concluded the comprehensive high school was sound other than these recommendations, and its existence remained much the same for the following two decades (Mirel, 2006).

The early 1980s brought changes to education. *A Nation at Risk* (ANR) was released in 1983 as a report on the American education system (National Commission on Excellence in Education, 1983). This criticism of the educational system suggested more rigorous coursework that was measurable and stricter requirements for a high school diploma (Park, 2004). These requirements included four courses in English, three courses in math, three in science, three in social studies, and one-half course in computer science.

ANR was the work of an 18-member commission created by then-secretary of education Terrell Bell (Ansary, 2007). The committee was formed to investigate correlations between the country's economic decline and its educational system. However, members of the committee later shared "they never set out to undertake an objective inquiry into the state of the nation's schools" (Kamenetz, 2018, para. 9).

Members of the committee already had a negative opinion of the nation's schools when they began their work. They looked for facts that would suit this narrative and used inflammatory language to bring more attention to the report (Kamenetz, 2018). When it was released, it had exactly the effect for which they had hoped, which was to convince the American people that public education was grossly failing students. The effects of ANR were felt into the 1990s. Many states adopted academic standards that all public schools were expected to adhere to for academic instruction. This led to an increased emphasis on high stakes standardized testing to measure these standards (Edvocate, n.d.).

However, there was criticism and contradiction to this report. Admiral James Watkins, who was then the secretary of energy, commissioned Sandia National Laboratories in 1990 to research if this decline could be documented with data (Ansary, 2007; Kamenetz, 2018). Their report, available in 1991, found that while there were some pieces of ANR that were accurate, students were actually performing the same as they had been or were slightly improving in almost every measure (Kamenetz, 2018). This was eight years after ANR was released, but ANR was still being used as an alarm for failing schools.

The Sandia Report as it came to be known was never published by the government (Ansary, 2007). It was sent into peer review and never became available after that, but it prompted multiple investigations (Miller, 1992). One investigation pertained to the how the

report was handled by officials at Sandia National Laboratories while another focused on how much of the Department of Energy's budget was being spent on educational research.

Furthermore, the analysts who wrote the report were told they were no longer allowed to work on anything related to education, and several expressed concern for the future of their careers (Miller, 1992). The report, which was finally published by the *Journal of Education Research* in 1993, is often used as a case study for censorship issues (Kamenetz, 2018).

Grades

Grades have been accepted as a necessary and critical component of the educational process (Schneider & Hutt, 2013); and as a core responsibility of educators (Adom et al., 2020). Anderson (2018) defines grades as “a position on a continuum of quality, proficiency, intensity, or value” (p. 3). Brookhart et al. (2016) refer to grades as symbols and note these are assigned to individual pieces of work or serve as composite measurements of a body of work.

Origin of Grades

American colleges provided the first evidence of assigning grades to students. Ezra Stiles, president of Yale University, noted in his 1785 diary, the rankings of 58 students who had taken an examination (Durm, 1993). The rankings were based on four categories: Optimi, second Optimi, Ineriores, and Pejores (Durm, 1993; Schneider & Hutt, 2013). These Latin terms were used to rank the quality of how well the student publicly expressed knowledge (Schneider & Hutt, 2013).

From this ranking, Yale developed a marking scale of four. Durm (1993) posits this is the genesis of the 4.0 scale currently still in use by many high schools, colleges, and universities. In 1813, William and Mary College used a four-category scale to classify students. These categories were not named but labeled by number classification (Durm, 1993).

Harvard University introduced a 20-point numerical scale in 1830 that was utilized for a rhetoric exam and would introduce a 100-point scale in 1837 (Durm, 1993). In 1877, Harvard began to classify students into six categories based on the 100-point scale introduced in 1837. Those categories are like modern grading scales: Division 1 was 90 or greater; Division 2 was 75 to 89; Division 3 included 60 to 74; Division 4 was 50 to 59; Division 5 encompassed 40 to 49; and Division 6 was anything below a 40. In 1884 they would reduce this to five categories without the 100-point scale (Durm, 1993).

In 1886, Harvard reported that categories one, two, and three were considered passing with distinction, category four was passing without distinction, and category five was considered failure to pass (Durm, 1993). However, in 1897, Mount Holyoke introduced the system that is mostly likely the foundation for modern grading scale and practices. This system was a combination of descriptive adjectives, such as excellent, fair, or failed, letters (A, B, C, D, F), and percentages based on a 100-point scale (Durm, 1993).

Early Grading

Prior to large-scale compulsory schooling, reporting of grades was traditionally a one-to-one conversation between the teacher and the families (Schneider & Hutt, 2013). Teachers visited the home to orally inform parents of the progress of the student (Guskey, 2015). Grading simulated the European model during this time and focused on competition among students and ranking members of the class (Schneider & Hutt, 2013).

In the late 1800s, increased numbers of students created a need for some form of grouping to reduce class size, so students began to be divided into age-based groups (Guskey, 2015). Grades were a relatively accepted practice at this time but were still subjective and represented in a variety of ways that changed frequently (Schneider & Hutt, 2013). Academic

learning was a means to the moral education of students at this time, which meant students would learn to read so they could read the Bible and “thus save one’s soul” (Vatterott, 2015, p. 7).

Increasing numbers of schools and students, in conjunction with the increasing mobility of families during the second period of the Industrial Revolution, created the need for a more standardized grading system. (Schneider & Hutt, 2013). The standardization of agriculture and industry was a model for the standardization of education. The considerable number of students attending school no longer allowed the teacher to visit the student home and share the progress of the student, so a stream-lined, time-conscious standardized approach was developed. By the start of the twentieth century as the number of students attending schools increased, grading evolved to become an external communication among “teachers, administrators, parents, college admissions officers and employers” (Schneider & Hutt, 2013, p. 3).

Purpose of Grades

Teachers identified six core reasons for grading:

- To document student and teacher progress
- To provide feedback to the student and family, and the teacher
- To inform instructional decisions
- To motivate students
- To punish students
- To sort students (Wormeli, 2006, p. 102)

Wormeli (2006) notes the first three are useful and helpful in advancing student learning; however, the last three dilute the purpose of the grade and are detrimental to student learning.

Motivating, punishing, or sorting students with grades can emphasize compliance and not learning.

The purpose of grades can also differ by grade band. In elementary grades, teachers are more likely to agree with the statements that students can learn without grade, and teachers can teach without grades (Guskey, 2009). Secondary teachers are less likely to agree with these statements, and they also are more likely to include homework assignments in the cumulative grade calculation. Cooper et al. (2006) noted this discrepancy may be the result of homework being more common in secondary grades. Overall, secondary teachers are more likely to adhere to more traditional views of grading (Guskey, 2009).

Grades can be useful when they are responsive to student needs and improve teacher practice (Anderson, 2018). When grades are analyzed through the lens of what content needs to be retaught or strengthened and what content has been mastered, they become a source of data to inform instruction. If teachers are not using them as part of a reflective practice, they become less useful for the teacher and do not impact future instruction.

Guskey (2006) proposes three types of grading criteria: Product criteria, process criteria, and progress criteria. Product criteria assesses what students know at a particular point in time; process criteria include the journey to the assessment, including homework, quizzes, compliance, attendance, and class participation; and progress criteria uses a growth model rather than a singular point in time. Of these, product criteria are the only ones that should be used as a source of data to drive instruction.

Homework

Homework is a collective term used to define any work that is expected to be done outside of regular school hours (Rosário et al., 2018). There are multiple purposes for

homework: to practice a recently learned skill, to recall a previously learned skill, to prepare for upcoming learning, or to extend learning (Rosário et al., 2018). Non-academic factors also influence homework. Its purpose may include to fulfill district or school expectations, to punish unwanted behaviors, or to communicate academic progress to parents.

Homework has been identified as a characteristic of schools and students who are able to reach achievement goals (Epstein & Van Voorhis, 2001). The additional time spent engaging with the academic skill during homework leads to better performance (Epstein & Van Voorhis, 2001), and homework also helps the development of non-academic skills, such as time management and self-direction (Keane & Heinz, 2019).

Conversely, homework can also negatively impact students and be biased toward socio-economically disadvantaged students. Kohn (2006) states there is “a direct relationship between how much time high school students spent on homework and the levels of anxiety, depression, anger, and other mood disturbances they experienced (p. 11). Socio-economically disadvantaged students may encounter difficulties completing homework due to inadequate home support or other obligations that prevent the student from completing the work (Keane & Heinz, 2019; Kohn, 2006).

Homework can also be a source of stress for teachers. Geving (2007) studied the impact on teachers of several different categories of student behaviors: harming school property, hostility toward other students, not being prepared for class, aggression toward the teacher, inattentiveness in class, lack of effort, hyperactivity, lack of interest, being noisy in class, and breaking school rules.

Geving (2007) observed moderate support in the data that lack of effort, aggression toward the teacher, and noisiness correlated with teacher stress; however, lack of effort was the

only category significantly correlated to teacher stress. This included completing homework on time. The author suggested motivation to complete work could come in the form of rewards for completing work and providing work that is interesting to the students.

Several researchers theorized that for homework to be impact learning it should be connected to the summative learning goal, high quality, and differentiated for student's needs (Epstein & Van Voorhis, 2001; Keane & Heinz, 2019; Kohn, 2006; Rosário et al., 2018). Students are less likely to complete homework if they feel it is busy work that does not have implications for the final learning goal (Vatterott, 2015). Teachers may assign less challenging homework for lower-ability students, and more engaging tasks for higher-ability students (Epstein & Van Voorhis, 2001). Other barriers to completing homework include unclear directions and/or expectations, lack of teacher feedback, and lack of support at home (Keane & Heinz, 2019).

For students with lower abilities, lack of differentiated homework can lead to frustration and incompleteness for a variety of reasons: lack of understanding, not having the correct resources, and lack of growth in their learning (Epstein & Van Voorhis, 2001; Keane & Heinz, 2019). Keith found homework improved grades for students who completed homework (1982; as cited in Epstein & Van Voorhis, 2001). However, Epstein and Van Voorhis (2001) found that low-ability students would spend 10 more hours per week homework compared to zero hours on homework for high-ability students to have "as good" (p. 184) report cards. This was true after all other factors, such as race, curricular track, etc., were controlled (Epstein & Van Voorhis, 2001).

Assessments

Formative and summative assessments are intended to measure student knowledge of specific standards or skills; however, the averaging of multiple assessments dilutes the intended message of the grade (Marzano, 2017; Wormeli, 2012). Averaging grades refers to the process of finding the mean of x number of scores (Khan, n.d.). This process does not provide the teacher, student, or stakeholders with any information about a specific standard or skill.

Averaging multiple assessments can misrepresent student proficiency, but individual assessments can also skew the measurement. Assessments often include items from multiple standards, so the final grade does not indicate which standards are mastered or to what degree they are mastered. Marzano (2017) argued that shifting to marking individual standards instead of an olio of questions would shift the perspective from one of assessment to one of measurement.

Teachers often model their assessment design after what they have experienced as students, and they often receive little training in assessments or grading (Guskey, 2006; Link & Guskey, 2019). Allen and Lambating (2001) theorize the shift for assessment practices should begin in education classes at the collegiate level. They believe it must be modeled before students adopt meaningful practices. Allen and Lambating also propose required measurement courses that address effective grading practices as part of the teacher education programs.

Formative Assessments. Formative assessments include a wide array of data gathering options. These are informal activities to gather instructions, such as conversations with students, questioning during instruction, homework, quizzes, and conferences with students (Burke, 2010). This feedback “is the heart and soul of formative assessment” (Burke, 2010, p. 21).

The data from these formative assessments is meant to guide the next steps in instruction and provide feedback to students on how to improve (Burke, 2010). This can be changes the student should make and/or shifts in the teacher's instruction (Burke, 2010; Films Media Group, 2008; Harlen & James, 1997). These instructional moves are meant to support the learners and move them toward a set learning goal either set by the teacher or themselves. (Films Media Group, 2008).

Harlen and James (1997) believe reliability is not essential to formative assessments because the information should only be used to inform instruction, and Kaftan et al. (2006) argue that daily work/homework and quizzes are used to gather points for grades rather than for feedback exchange between the teacher and the student. Many formative assessments would not be graded at all but would only provide corrective feedback to the students. Kaftan et al. (2006) posited that the success of formative assessments without a grade and only feedback required teachers to build relationships with students to earn their trust by listening to the information they provided in the assessment and acting upon it.

Summative Assessments. While formative assessments are meant to inform next instructional moves and provide feedback to students, summative assessments are intended to measure the cumulative knowledge of the student in terms of the content of the assessment (Ahmed et al., 2019). These assessments indicate to the teacher the students' success or failure on the learning goal(s) or unit of material (Ahmed et al., 2019; Burke, 2010).

Summative assessments do not have the variety of forms compared to formative assessments. Types of summative assessments include teacher-created exams, standardized tests, curriculum-created exams, or a final project (Yale Poorvu Center for Teaching and Learning, 2021). Burke (2010) notes there is typically no feedback provided on a summative assessment

other than a numerical grade. Because of this, summative grades are considered feedback for the teacher's instruction for future students more so than to support the current learners (Ahmed et al., 2019; Burke, 2010).

Standardized Assessments. Tennessee requires End-of-Course assessments for high school students in English I, English II, Algebra I, Algebra II, Geometry, Integrated Math I, Integrated Math II, Integrated Math III, Biology, U.S. History, and Geography. These tests are “administered in accordance with the Every Student Succeeds Act (ESSA) of 2015 and T.C.A. § 49-1-602” (Tennessee Department of Education, n.d.-c, TCAP End-of-Course Section, para. Federal and State Law). These assessments are a specific type of summative assessment.

State board policy requires scores from the End-of-Course exams to be used as a part of the student's final grade for that course (Tennessee Department of Education, n.d.-c, TCAP End-of-Course Section). State law requires the results of an End-of-Course exam to be calculated at a minimum of 15% of the final course grade if the results are available to schools more than five instructional days from the last day of school for students (Mangrum, 2022; Tennessee Department of Education, n.d.-c, TCAP End-of-Course Section).

In addition to End-of-Course exams, Tennessee students are required to take a college entrance exam, either the ACT or SAT, during eleventh grade to receive a regular or honors diploma (Tennessee Department of Education, n.d.-c, ACT and SAT Section). These scores are not calculated as part of a course grade.

Measurement

Teachers measure students in multiple ways: summative assessments, written responses, homework, etc. These measurements should be representative of the full body of knowledge of the student at that moment in time (Guskey, 2015; Schimmer, 2016), but teachers do not have

sufficient understanding of measurement and how to determine or interpret academic measures (Lambert, 1989). According to Guskey (2009), “few teachers receive any formal training on grading and reporting” (p.2). Many teachers lack the preparation and understanding necessary to calculate a measurement that carries such an impact on students. For teachers to accurately communicate this to parents and students, grades must have validity and reliability.

One function of grades is to communicate a child’s academic progress to parents and/or stakeholders. While parents express greater confidence in a numerical grade (Guskey & Bailey, 2010), there is a not a clear understanding of what a grade means whether it is a letter grade or a numerical grade (Tomlinson, 2005). Additionally, parents are unsure of what other grading scales, such as narrative grading scales, represent. These grading scales use terms such as mastery or approaching to describe the level of mastery.

Validity

When a measurement is defined as valid, it means that it accurately measured what it was intended to measure (Chiekem, 2015; Fitzner, 2007). Grades should have validity; however, single-task assignments are most often considered to be unreliable (Anderson, 2018). This creates questions about the validity of the cumulative measurement. In education, the cumulative measurement is the grade at the end of the marking period, semester, or year.

Brookhart et al. (2016) found researchers have questioned the validity of grades for over 130 years. The earliest study by Edgeworth in 1888 identified differences among the graders as a reason for concerns about validity (Guskey, 2015). A 1927 study by Bolton found a greater variability in grading when the papers were a poorer quality (as cited by Brookhart, 2016); however, Ells found in his 1930 study there was a higher consistency in grading when the paper

was of the poorest quality (as cited by Brookhart, 2016). Additional studies have continued to support these findings (Anderson, 2018; Brookhart, 1993; McMillan & Nash, 2000).

Reliability

Because grades impact self-concept of students and the opportunities available to them based upon these grades, it is important grades are reliable (Chen & Bonner, 2017; Chiekem, 2015; Griffin & Townsley, 2021; McMillan & Workman, 1998; Stiggins, 2002). Reliability in grading is defined as “the consistency of scores across replications” (Miller, 2019, para. 1). If a student completed the same assessment multiple times with the same outcome, the assessment would have a higher reliability than an assessment where the outcomes were different.

The reliability of grades and grading practices have been scrutinized for over a century (Brookhart et al., 2016). Haladyna (2019) stated that a cumulative grade consisting of only a mid-term and a final exam has low reliability because there is low opportunity for replication. Providing other opportunities for students to replicate their level of knowledge would increase the reliability of the grade. Quizzes, written assignments, and other activities or accomplishments are examples of this.

Starch and Elliott (1912) studied the reliability of grades in the early 1900s high school English courses and found wide variations and “the utter absence of standards in the assignment of values” (p. 442). The initial study researched high school English essay grades and tasked English teachers with grading two essays written by two different students (student A and student B) at the end of the course. Copies were sent to 200 first-year English teachers, and 147 were usable when returned. The researchers found there was a wide range in the grading (Starch & Elliot, 1912). Although teachers usually predict a range of “at most 10 points” (Starch &

Elliot, 1912, p. 454), the authors noted a range of “as large as 35 or 40 points” (p. 454) was possible when multiple teachers graded the same paper.

Critics of the study stated the discrepancy was expected because of the subjective factors involved in marking an essay (Starch & Elliott, 1913). Starch and Elliott conducted a second study focused on math to test this theory. The math study asked teachers to grade a high school geometry paper completed by a student as the final exam for the class (Starch & Elliott, 1913). The set of questions and a copy of the student paper were sent to 180 high schools, and 140 papers were returned. Of the 140, twelve were discarded. The schools had different minimum passing scores, but the data was adjusted for this. This second study not only reinforced the findings of the first study, but the data also pointed to an even wider range of grade distributions for the geometry paper compared to the previous study.

A math paper was used because it was assumed by Starch and Elliott that math was assessed with a precision that could not be found in an English course. However, the authors stated this was illogical thinking:

A little analysis, however, will show the absurdity of assuming greater precision in evaluating a mathematical paper than in evaluating a language or any other kind of paper. While it is true that there can be no difference of opinion as to the correctness of a demonstration, yet there are countless ways in which the demonstration may be worked out, involving the succession of the steps, the use of theorems and definitions, the neatness of the drawings, and most of all the relative value of each particular demonstration or definition in the evaluation of the paper as a whole. Obviously the complication of factors is as intricate in one sort of paper as in another. (Starch & Elliott, 1913, p. 257-258)

Because 22 of the teachers in the original study did not give Student B a passing grade, the researchers noted the impact this could have had on that student, such as failure of the course or retention. (Starch & Elliot, 1912). From these studies, Starch and Elliott (1912, 1913) reasoned that student grade-level promotion or retention moderately relies on the subjective analysis of the work by the teacher.

Brimi (2011) sought to recreate century-old Starch and Elliott's work with two minor differences: one, the grading scale across the state in which the research was conducted had been streamlined; and two, each of the 90 teachers who participated had been trained in a specific grading system for writing. Teachers were trained for a year using this tool with a final follow-up session prior to the paper analysis. The persuasive paper was written by an anonymous eleventh-grade student who volunteered her paper to be used for the study.

The participants first independently created a rubric using the writing grading system. They were specifically asked not to consult with other teachers during this process. Once the rubric was completed, they were asked to score the paper. Of the 90 participants, 89 participated (Brimi, 2011). Of those, three scores had to be discarded for different reasons that made the data unusable. Seventy-three of the remaining 86 papers had followed the directions and provided usable data.

The research validated Starch and Elliott's findings from a century earlier: Teachers who had engaged in the same trainings scored the same paper very differently (Brimi, 2011). The range of lowest score to highest score was 46 points, and 30 different scores were assigned to the paper. The most common scores were either the minimum or maximum for a letter grade cut-off, such as an 84%, which was the maximum score for a "C" at that time. When numerical grades

were equated to letter grades, there were ten scores of “A”, 18 scores of “B”, 30 scores of “C”, nine scores of “D”, and six graded the paper as an “F”.

Brimi (2011) noted in his research that the topic of the paper could be controversial but went on to note many of the teachers did not generally comply with the tool. He observed that many teachers offered few to no feedback remarks on the paper and comingled a prior tool into the grading process. The author posited in summary that this type and level of disparity could lead loss of scholarships, denial of university admission, teacher-shopping in schools, and loss of teacher autonomy in instruction as schools and districts try to decrease subjectivity in writing assessment.

Equating of Scores

If a single version of an assessment is given, all students will answer the same questions so their scores could be compared directly; however, some assessments have multiple versions. Most large-scale assessments, such as state standardized tests, have multiple versions of the assessment, and some teachers will use multiple versions of classroom assessments to discourage cheating (Ryan, 2018; Stauffer, 2022).

When there are multiple versions of an assessment, equating is conducted to create comparable scores among the different versions. This impacts the validity, or interpretation, of the test scores. It is needed to adjust for the differences in difficulty levels in questions among the different versions (Issayeva, 2021). Equating is stronger than linking because it claims that scores from multiple versions of the same assessment have the same meaning whereas linking only posits there is an association between the scores (Ryan, 2018).

When different forms of an assessment are equated, it can be said they measure the same content and require the same cognitive processes; thus, the same conclusions can be made of the

scores from different versions of the assessment (Ryan, 2018). Equating also takes place on standardized testing from one year to the next. An End-of-Course assessment is equated among the different versions in one testing year, and they are also equated to prior years' tests (Ryan, 2018). This allows scores to be compared between cohorts.

There are two ways to equate assessments: Classical test theory (CTT) and item response theory (IRT). IRT is considered a stronger method for equating than CTT; however, CTT works better with smaller sample sizes (between 100 and 300) and is used more often (Issayeva, 2021; *The Elements of Mental Tests*, 2018).

Both theories are used to convert a raw score to an adjusted measurement that can then be compared to different versions of the assessment whether that be different forms within the same year or the assessments across multiple years (Donovan, 2017). According to Wright (1977), this is necessary because raw scores will never be re-created exactly and are considered temporary descriptions.

Often equating is done after the assessments are given and the practitioner has the full data set. It can also be done prior to the assessment, such as when there is not a set window to administer the assessment, but this makes the equating process more challenging because the data is continuously being gathered (Issayeva, 2021).

Ryan (2018) states, "It is critical to note that the process of equating test forms begins with the very design and construction of those test forms" (p. 11). Psychometrics is the field of designing, scoring, and analyzing data from assessments, which is often an area of study neglected in pre-service teacher education programs (Stiggins, 1993; Torpey, 2011). Most practicing psychometricians have a Ph.D., and the minimum education is a Master's degree. These degrees are often in the areas of educational measurement, quantitative psychology, or

statistics (Torpey, 2011). Classroom teachers often do not have this level of expertise but create and/or use teacher-designed assessments in the classroom, and this can contribute to assessment grades that are not valid or reliable being included in the final course grade (Lynch, 2016).

Square Root and Cube Root Calculations

Post-equating assessments can take copious amounts of time after they are completed. To return scores to a student or school system more quickly, a square root curve (sometimes referred to as a Texas curve) or cube root curve is sometimes used to convert raw scores to scale scores. Tennessee uses a cube root curve for the scale scores that are then used as part of the final course grade for certain high school courses (Spears, 2015).

To calculate scale score from a raw score, the raw score is first converted to a conventional percent grade by dividing the number correct by the total number of items on the assessment and multiplying this decimal by 100. This number is then calculated to the power of $\frac{1}{3}$ and, finally, multiplied by 21.5443, which is a constant term to create a maximum calculation of 100 (Tennessee Education Association, n.d.).

This method can be skewed to inflate lower raw scores into higher scale scores. For example, a raw score that converts to 83% would have a cube root score of 93.98%. A raw score that converts to a 43% would have a cube root score of 75.48. The first score was only raised 10.98 percentage points while the second score increased by 32.48 percentage points. Additionally, the second score has moved from a failing score under Tennessee's grading scale (below 60) to a grade of "C" which is passing.

Grading Systems

Several central grading systems serve as the foundation of grading practices in the United States, and there is no national expectation or mandate for these (International Affairs Office, n.d.). These methods are the central systems to communicate achievement to students and stakeholders (Mcbride & Kahle, 2018). Traditional grading systems are often associated with letter grades assigned based on a point scale, and criterion-referenced systems are traditionally the base for these (International Affairs Office, 2008). Other grading systems include norm-referenced and standards-based.

Criterion-Referenced Grading System

Criterion-referenced grading systems use assignments and assessments using criterion-referenced design to compare student knowledge against an accepted fixed or pre-established standard (International Affairs Office, 2008; Renaissance, n.d.). Students are not compared to other students in the group but are measured according to the standard based on their individual performance (Lynch, 2021).

There is no limit to the number of students who can be assigned a particular numerical or letter grade because the grades reflect what the student has learned (Guskey, 2015), and the grading scale is set before the assessment is given (Hall, 2013). Pui et al. (2021) found that information from criterion-referenced assessments is the most useful in understanding student abilities and needs, the most transparent for students to understand, and the most helpful in responsive teaching.

Norm-Referenced Grading System

Norm-referenced grading assumes all students within the group have comparable ability levels (International Affairs Office, 2008). This system, which is associated with grading on a

curve, is pre-established with percentage ranges that assume a five-letter system of A, B, C, D, and F (Öztürk Gübes, 2021). The A and F ranges have an equal number of scores; B and D also have the same number of scores; and C contains the largest number of scores. An example of this would be 10% of scores would be assigned an A or F, 25% of scores would be assigned a B or D, and the remaining 30% would receive a C.

Scores can be normed within the classroom group, but they can also be expanded to the district, state, and national levels. They can also be delineated by age, language status, socioeconomic level, race, or ethnicity (Renaissance, n.d.). Practitioners use norm-referenced grading to communicate percentile ranking against others in the group, but it cannot measure specific levels of knowledge.

Norm-referenced grades can be misleading. Öztürk Gübes (2021) found weighting grades or categories of grades in a norm-referenced grading system can significantly impact student grades when two or more scores with different standard deviations are combined. The score with the larger standard deviation would have a greater impact on the final grade. The study revealed standardization of scores is the only way to avoid skewing the composite score.

Another complication for normative grading practices is the creation of a risk-aversion culture. This system determines student success relative to other students, thus students are trying to out-perform their peers rather than focusing on proficiency in the content (Burleigh & Meegan, 2016). This leads to less cooperative interactions among students and less sharing of knowledge in classes. Burleigh and Meegan (2016) also theorize this impacts social and emotional skills and fosters less resiliency in individuals.

Standards-Based Grading System

The standards-based grading system is based upon the standards assigned to a specific grade level and/or content area and does not include any external factors (Schimmer, 2016; Wormeli, 2006). Teachers provide feedback either as a narrative or a score based on a pre-established rubric or scale for individual standards rather than measuring multiple standards at once and averaging a numerical grade for them (Bouchrika, 2021; Iamarino, 2014). In this system, teachers and students are both continuously aware of the level of proficiency for the stated standards.

Standards-based teaching emerged in the early 1990s as the federal government sought to reform education (Hurst et al., 2003). One of the four categories of reform was standards, assessment, and accountability. Standards-based grading (SBG) emerged from this as a way for teachers to provide differentiated instruction and be responsive to student needs (McBride & Kahle, 2018).

Guskey et al. (2020) argues in favor of standards-based grading because it separates the cognitive (academics) from the non-cognitive, such as behavior or homework completion, and reporting these individually. Feldman (2019) asserts that awarding points that are included in the grade for behaviors the teacher prefers “imposes on students a culturally specific definition of appropriate conduct” that may be biased against him (p. 54). Feldman (2019) posits that excluding behaviors in grading lessens potential teacher bias toward the student.

Percentage Grading System

High school enrollment grew very quickly after compulsory education laws were passed (Guskey, 2013; Gutek, 1991). With this increase, narrative reports to parents became a challenge, so many high school teachers began to streamline the process by using more abbreviated

information, such as percentages (Guskey, 2013). This allowed teachers to communicate more quickly with students and teachers.

Percentage grades became a fixture of grading during the early 1990s with the increased use of online or computerized grading programs (Guskey, 2013). These programs allowed teachers to input grades, and the program would complete the necessary calculations, including weighting certain grades or categories higher than others. Criterion- and norm-referenced assessments may be included in a percentage grading system.

In this system, there are two measurements being used: percentage correct, which is found by dividing the number of items correct by the total number of items, then multiplying this decimal by 100 to determine the percentage correct (Witte & Witte, 2017), and the mean of all scores considered for the cumulative grade. For a basic mean, all scores are added together, then divided by the total number of scores; however, some teachers will weight some categories higher than others, such as exams. To determine these grades, the mean is found for each category, then those values are multiplied by the decimal form of the weight (60% = 0.6). Finally, each weighted value is added to determine the final grade. Table 1 shows an example of this measurement.

Table 1

Percentage Score Example

Category	Homework	Quizzes	Exams
Percentage Scores	75%, 70%, 63%, 78%, 88%, 95%, 90%, 100%, 100%	72%, 86%, 65%, 90%, 97%,	76%, 82%, 91%

Category	Homework	Quizzes	Exams
Percentage Mean for Category	84.33%	82%	83%
Weighted Percentage	20%	20%	60%
Final Category Weight	16.87%	16.4%	49.8%
Final Score	83.07%		

Grading Scales

Grading scales based on a 100-point system are problematic because the accepted range for a passing grade is smaller than the range of failing. A 100-point grading scale that uses the scale shown in Table 2 (Tennessee S. Rep. No. 388, 2022), demonstrates this narrow distribution. Guskey (2013) posits that a larger range for a failing grade implies that failing grades can be defined in better detail than passing grades. A teacher can choose 11 ways to define an A, but there are 60 discernible levels to select from within the F grade range.

Table 2

State of Tennessee Grading Scale for Grades 9-12

Letter Grade	Numerical Range
A	90-100
B	80-89
C	70-79
D	60-69
F	0-59

While this can be perceived as being more precise, the actual precision comes from the accuracy of the measurement tool, such as an assessment or assignment.

The 100-point scale can create more opportunity for student achievement to be misclassified compared to a scale with lesser categories, such as only marking a letter grade with no numerical value attached:

A student is statistically much more likely to be misclassified as performing at the 85-percent level when his true achievement is at the 90-percent level (a difference of five percentage categories) than he is of being misclassified as scoring at an *Average* level when his true achievement is at an *Excellent* level. (Guskey, 2013, p. 3).

Grading Practices

Grading practices encompass the ways teachers report and use single-task grades, such as an essay or unit test, and cumulative grades at the end of the grading periods (Guskey & Brookhart, 2019). While research related to grades and grading practices extends over 100 years, the quantity of studies is limiting. Brookhart et al. (2016) completed a meta-study and identified 35 studies since 1994 that incorporated teacher perception of grading and grading practices. Brookhart et al. (2016) found:

- There was a significant diversity in grading practices among teachers.
- Teachers had little awareness of district policies about grading.
- Ability and effort should be included in grading practices.
- Teacher experiences as a student and internalized values about grades influence teacher grading practices.
- Non-achievement factors are included in grading practices to help students get the highest possible grade.

- Teachers incorporate student personality and behavior into grading practices.
- Female teachers emphasized behavior in grading practices more than male teachers.

Teachers do not intentionally choose grading practices that do not accurately measure learning. Rather, teachers often use a random approach that is not mathematically sound due to concerns for students, including motivation and social consequences (Guskey, 2006). Guskey (2009) stated, “students are likely to face strikingly different grading practices as they move from class to class within the same school” (p. 11). Lack of clear grading policies and understanding of accurate measurement practices contradict the true meaning and understanding of grades by teachers, students, and stakeholders.

McMillan and Nash (2000) suggested that teachers adopted these approaches because they were trying to find a practice that would help determine a successful grade for the students. While this practice may protect students in the short term, it creates a false sense of success in the long term for students who may later experience failure and not understand why.

Preservice teachers who have not had adequate coursework in grading can be unaware of the appropriate methods and ethical responsibilities of grading (Bonner & Chen, 2009). Their research found the elementary and secondary preservice teacher participants both opposed traditional grades but supported using grades to manage behavior; however, the secondary preservice teachers were in favor of using grades to punish a single student or groups of students.

Bonner and Chen (2009) did find these preservice teachers’ grading positions did shift after completing coursework in classroom assessment. The authors suggested this could be because they had learned more about fairness and validity related to grading. They did note,

however, this tempering of opinion may not last if there is not additional professional development related to this.

Grading as Moral Responsibility

There is also a moral perspective for teachers and grading. Humans adopt morals to discern between what they believe to be right and/or wrong, and morals act as the binder among societies (McCombs School of Business, n.d.). Declaring an action as fair or defining an act as unacceptable are examples of morals.

Zoeckler (2007) posited that grading has a deep moral scope because it requires answering what is fair and what is good for students. His study focused on teachers in an English department in a rural upstate New York high school. They were asked questions that inquired about the mathematical calculations of grades, such as weighting categories, and perceptions of fairness. The research showed teachers used judgements of student character as a factor in grading.

Teachers choose or alter grading practices to achieve a certain outcome. Zoeckler (2007) found that the fixed-time intervals (mid-term, end-of-grading period) could falsely represent the student grade because of this. Additionally, teachers sometimes manipulated the weighting of grades “to offset poor class averages” (Zoeckler, 2007, p. 92). One participant stated she looked for patterns in student work and would discard any grade – better or worse – if it did not fit the established pattern.

The study also found decisions about student character factored into decisions about grades. Although none of the teachers stated character should not be a factor in grading, all admitted to incorporating characteristics such as diligence, carelessness, and laziness into grades (Zoeckler, 2007). These are moral descriptors that were subjectively assigned by the teachers.

Participants indicated these were indirect effects that occurred when deciding to pass or fail a borderline student. Students judged to have good character were given a passing grade whereas a student thought to be of bad character failed (Zoeckler, 2007).

Mindset, Self-Concept, and Mental Health

Mindset and self-concept impact student success (Dweck, 2016; Hattie, 2009). A student envisions themselves as a particular type of student, whether successful or not, and can develop a specific mindset around this vision. Grades affect the mental health of teachers and students: Both feel pressure to meet certain expectations around grades (Feldman, 2020).

Mindset. Dweck (2016) posits there are two mindsets: fixed or growth. A person with a fixed mindset is discouraged by failure and believes intelligence to be a rigid trait while a person with a growth mindset feels challenged in a positive way by failure and believes intelligence can be cultivated. Those with a fixed mindset are focused on the outcome, but a growth mindset does not need this external validation.

Transitions, such as moving from elementary school to middle/junior high or from there to high school, can often challenge mindset. Students will use their mental resources to protect their egos rather than risking what others think of them, so they often do not try new things, including academics, if they consider it to be a threat (Dweck, 2016; Yeager et al., 2019). Rather than risk failure or losing credibility among their peers, they will accept lower grades: “Perhaps students in unsupportive peer climates risked paying a social price for taking on intellectual challenges in front of peers who thought it undesirable to do so” (Yeager et al., 2019, p. 368).

Teachers also play a role in this process. If teachers have a fixed mindset about students, it can affect the learning experience for the student. Nottingham and Larsson (2022) describe two students entering a classroom for the first time. Student A is described as a young girl who has

been referred to by previous teachers as gifted and is intrinsically motivated to learn whereas Student B is a young boy who has been described by staff as being below average, a reluctant reader, and needing extrinsic motivation. If the teacher has a fixed mindset about these students – being more supportive of Student A or Student B - the learning environment is not neutral for either student.

Nottingham and Larsson (2022) explain a teacher with a fixed mindset is not beneficial for either student. If the teacher is favorably biased toward Student A, the young girl may grow to have a false sense of security about her abilities and then struggle when she is challenged in her learning. If the teacher is negatively biased toward Student B, he may believe he will never be successful at anything and stops trying to become a better reader. Both positions are an extraneous impact on students' grades.

For students and teachers who seem to be stuck in a fixed mindset, a growth mindset intervention could shift their position. Yeager et al. (2019) provided a growth mindset intervention for 6,320 ninth-grade students who were considered lower-achieving relative to their peers in the core areas of mathematics, science, English/language arts, and social studies. The results showed students who received the intervention were earning higher GPAs in these core subjects by the end of their ninth-grade year. Another effect of the intervention was increased enrollment in advanced mathematics (algebra II or higher) in tenth grade.

Teacher mindset affects learning, which can affect grades. Dweck (2016) states, “Fixed mindset teachers often think of themselves as finished products. Their role is simple to impart their knowledge” (p. 204). These teachers perceive learning as the student's responsibility. Growth mindset teachers love the process of learning and engage in this process along with their

students. If the students are not learning, these teachers are reflective about their role in that lack of learning.

Self-Concept. The beliefs one holds about themselves play an integral role in academic success, and grades impact this, especially for those with a fixed mindset. Students whose past experiences with grades have been positive have a positive self-concept in relationship to grades. Those who have had negative experiences or have often received failing or significantly lower grades tend to feel less intelligent and capable when comparing themselves to others (Dueck, 2020).

As the education system has shifted since ANR toward more standardized testing, many students have developed a sense of hopelessness because they have not been able to be successful on these assessments (Stiggins, 2002). When students perform poorly over time, the concept of failure is incorporated into their self-concept beliefs. Rather than thinking of a poor grade as a single experience, they begin to view themselves as incapable of success.

The relationships students have with teachers influence student self-concept. McFarland et al. (2016) found that how students perceived their teachers significantly impacted their self-concept. This was true for both boys and girls. When students had a close relationship with their teacher, they were more engaged in school. Research has indicated teachers include moral judgements in their grading practices, and this has implications for student self-concept.

Högberg et al. (2021) found the introduction of grades led to increased stress and anxiety and lower academic self-concept. The research also noted evidence of psychosomatic symptoms associated with grades that impacted quality of life. This effect was felt more strongly among girls. Giota and Gustafsson (2021) had similar outcomes in their research.

Mental Health. Grades can also impact the mental health of students and teachers.

Mental health can be the reason grades drop, but lower or failing grades themselves can also be the cause of poor mental health. Feldman (2020) states, “Relentless pressure to succeed, often measured by grades or GPA, can contribute to students being sleep-deprived, anxious, or even engaging in self-harm” (para. 4).

Determining grades can be stressful for teachers. Although grades are meant to reflect student learning, the assessment of that is reflective of the expertise of the teacher in not only the content area being assessed, but also the professional judgement of the teacher. When a student or parent disagrees with a grade, they are in essence questioning the professional knowledge of the teacher (Feldman, 2020).

Teachers often do not feel confident in their grading practices because teachers receive little to no training in grading during their pre-service experience or as practicing teachers through professional development. Thus, they will typically replicate their own experience as a student, use the same system as their colleagues, or use a hodge-podge system they create (Feldman, 2020). This lack of confidence can lead to additional stress for the teacher.

Academic and Non-Academic Factors in Grading

Although grades are assumed to be accurate representations of student learning, they are often composed of several factors that are not related to academics. As early as 1936, research showed the influence of non-cognitive components on grades (Sobel, 1936, as cited in Brookhart et al., 2016). Students in this study had low test scores but high overall grades. The study found the students had high marks for “penmanship, attendance, punctuality, and effort” (Brookhart et al., 2016, p. 822) and were rated highly for qualities such as perseverance and cooperation.

This is commonly referred to today as grade inflation. Static grade inflation is a singular, point-in-time overstatement of grades, while dynamic grade inflation occurs when there is a change between grades and achievement (Tyner & Gershenson, 2020). This is not a point-in-time event but happens over time. Differential grade inflation takes place when there is a discrepancy among schools or students in their grades and achievement.

One identified purpose of grading is to communicate academic achievement; however, this purpose is lost if the information in that communication is faulty. Stanley and Baines (2001) described a grade as “the cornerstone for communication” (p. 227) among students and other stakeholders and posited that if this cornerstone fails then the entire system could eventually fail.

In their study, Tyner and Gershenson (2020) stated that grade inflation is not well defined as a concept, but it does exist and impact students. They analyzed course grade and end-of-course standardized tests and determined dynamic grade inflation was taking place in schools where more advantaged students were being served; however, this was closing an existing grade inflation gap between these students and those students in low-advantage schools where static grade inflation was the norm. This study showed the different elements of grade inflation and how they each misrepresent student knowledge and understanding.

Grades are often a mixture of achievement and non-achievement measures, such as combining assessment scores and effort (Cross & Frary, 1996; Link & Guskey, 2019), and this conflating of achievement with other non-achievement factors damages the validity of the grade. Griffin and Townsley (2021) found in their study that 61.5% of the final math scores they analyzed had been inflated when the categories of homework and employability, defined in the research as skills such as punctuality, participation, and citizenship, were included. It was also noted this inflation in some cases was 5% or more, which could increase a student’s letter grade

by half or increase it to the next letter grade (Griffin & Townsley, 2021). Additionally, they discovered 10% of the students would have failed the course if the homework grades and employability scores were removed from the final score.

Stiggins et al. (1989) recommended achievement be the only factor considered in grading, and measurement specialists confirm this recommendation. When an elementary teacher combines an assessment grade covering addition, a worksheet categorizing place value, and participation during centers, the meaning of the grade is blurred (McBride & Kahle, 2018).

Research has shown the “hodgepodge of factors” (McMillan, 2001, p. 28) used to calculate a grade can be broken down into four categories: academic achievement, academic enablers (effort, participation, improvement), external benchmarks, and using extra credit or teacher-determined decisions for borderline cases. This approach creates a lack of clarity for the meaning of the grade and distorts the proficiency of the student (Guskey, 2006). This is in contrast with the recommended best practice by measurement specialists of calculating grades based solely on achievement (Cross & Frary, 1996).

Grades are a form of currency in many classrooms. Grades are awarded to students as payment for a combination of an acceptable amount of effort, a positive attitude, punctuality, attendance, participation, and compliance, in addition to the cognitive progress in learning (Brookhart et al., 2016). This practice leads to manipulation of the grading system by both teachers and students.

Expanding on the characteristics mentioned by Brookhart et al. (2016), Westphal et al. (2021) used the five-factor model used to describe typical behavior of a person, also called the Big Five personality traits, to research grading discrepancies in secondary schools. The Big Five personality traits are openness, conscientiousness, extraversion, agreeableness, and emotional

stability (McCrae & Costa, 1997, as cited by Westphal et al., 2021). Their study examined the relationship between teacher-assigned grades and student personality considering discrepancies on standardized test performance.

McCrae and Costa determined there was a significant variance between teacher-assigned grades and standardized test performance. They found two of the Big Five traits to be the best predictors of the teacher-assigned grades: conscientiousness and agreeableness. They posit conscientiousness positively impacts the teacher-assigned grades because these students are more invested in classroom goals than less-conscientious students. This may also reflect the value system of the teacher.

Conversely, agreeability negatively affects student grades. Agreeable students tend to prefer situations where everyone benefits and there is no competition. Westphal et al. (2021) theorize agreeable students are negatively impacted because they view grades as a competition, which is problematic for them. Whether grades are positively or negatively impacted, including personality traits in grading skews the measurement of knowledge and understanding.

Grades as Motivators or De-motivators

While grades are an accepted and expected piece of education, grading can have multiple purposes. They are intended to communicate student knowledge to stakeholders, but teachers also use them as motivators to promote desired behaviors (Chiekem, 2015; Vatterott, 2015). These behaviors can be academic or non-academic.

Using grades as motivators does not increase the academic achievement of students. Grades can serve as de-motivators and may be viewed as “controlling” (Ryan and Deci, 2020, p. 6), which decreases autonomous, intrinsic motivation. Grades are potentially the only extrinsic

reward; however, teachers must promote specific norms or discourage certain behaviors (Andersen, 2018).

Grades as Punishment or Reinforcement

Grades can also be used as punishment for unwanted academic and non-academic behaviors or a reinforcement for desired behaviors (Chiekem, 2015; Vatterott, 2015). Guskey (2009) found this to be more common in secondary classrooms than elementary classrooms. Guskey (2009) found that secondary teachers were more likely to assign zeros for missing work and deduct points for late work or behavioral infractions. He proposes this is a form of control over students as grades can indirectly shape privileges or punishments that may be given in the home.

Hochweber et al. (2014) found that teachers used grades as a form of classroom management to establish control. Deducting points for lack of a name on a paper or giving extra points for attending an optional school function, act to leverage points for a wanted behavior. Hochweber et al. (2014) found this grading behavior to be more common in classrooms that were lower performing academically and attributed this to these classes being more demanding from a behavioral perspective.

Administrative and Parental Expectations

The Every Student Succeeds Act (ESSA) is the federal K-12 education law for the United States, and increasing graduation rates, while at an all-time high when it was passed in 2015, is a component of this act (U.S. Department of Education, n.d.). Increasing graduation rates is also part of the accountability model for school systems in Tennessee (Tennessee Department of Education, 2022). Students who do not graduate with their cohort affect this rate even if they

were to graduate the following year. This can create pressure on teachers to provide passing grades even if a student has not mastered the material.

If a student fails a course that is required for graduation, they would be required to retake the course. This may take place in the summer if available or the following school year. Repeating the course affects the funding for the school and could reduce the amount of money available per student (Gonzalez, 1995).

Washington Teachers' Union and EmpowerED found in their 2017 survey of District of Columbia teachers that 46.5% said "they felt pressured by a school administrator to pass or change grades for students who didn't meet expectations" (Johnson, 2018, para. 3). When narrowed down by grade bands, 60% of high school teachers reported feeling pressured. Teachers described this pressure as being the greatest in high poverty neighborhoods.

In addition to pressure from administration, teachers indicate parents also pressure them to change grades or decrease expectations for students. Lloyd (2019) speculates this is because of high schools not being as rigorous as they should, fewer spots in strong universities, and extreme competition for scholarships. Grades are described as the tollbooth by which students gain access to these opportunities. To parents, teachers are in essence the gatekeepers to their student's success.

Trauma

Learning can be a challenge for students who have been exposed to childhood traumas, also called Adverse Childhood Experiences (ACEs). Examples of an ACE include emotional, physical, or sexual abuse; physical neglect; parental divorce; drugs and/or mental illness in the home; and/or, having a household member in jail or prison (Iachini, 2016). The higher the number of ACEs, the higher level of trauma.

The Brain. Trauma can affect the development of the brain and its ability to process during critical development periods such as childhood (Porche et al., 2011). Learning begins in the brain stem, which is the lowest part of the brain (Sprenger, 2020). This part of the brain is responsible for different roles. Calculating responses to stimuli and acting as a filtering system are two of those (Johns Hopkins Medicine, n.d.; Sprenger, 2020). It is also responsible for focusing vision and facial expressions (Johns Hopkins Medicine, n.d.).

If information is somehow threatening to the individual, the brain stem activates the fight-or-flight response to the rest of the body. This response releases a flood of stress hormones that increase heart rate, tense muscles, quicken breathing, and/or produce sweat (Harvard Health Publishing, n.d.). This process happens very quickly, and people are typically not aware that it is happening until they feel the physical reactions, such as increased heart rate (Harvard Health Publishing, n.d.).

If information passes the brain stem, it moves to the limbic system. In this system, the amygdala acts as a second filter for information (Sprenger, 2020). It is also in charge of the emotions attached to the information and regulates memory (Johns Hopkins Medicine, n.d.). The hypothalamus, which is where the chemicals for the fight-or-flight response is stored and released from, is also located in this system. It controls body temperature, plays a role in memory and emotion, and controls hunger and thirst (Johns Hopkins Medicine, n.d.).

The destination for information is the frontal lobe. Executive function and working memory take place here (Arnsten, 2009). This part of the brain is where decision making, problem solving, and attention take place (Guy-Evans, 2021). The prefrontal cortex is part of the frontal lobe, and it is responsible for higher cognitive functioning. Once information reaches the

prefrontal cortex, it can be retrieved in future situations (Arnsten, 2009; Guy-Evans, 2021; Sprenger, 2020).

Impact of Trauma on Learning. For learning to be accessible in the future, it must find its way to the prefrontal cortex. Trauma can be a barrier to that pathway. Duplechain et al. (2008) researched the impact of trauma on students' reading achievement and found a statistically significant relationship between lower reading achievement and violence exposure. They determined this is because these students lack at least one of the following: motivation, concentration, or the ability to connect to personal experience. They hypothesized the most exposed students would be the most at-risk; however, the data showed moderately exposed students were more at-risk for lower reading achievement. They theorize this is because higher-exposure students are considered more at-risk by schools, and moderately exposed students do not exhibit the same academic struggles as higher-exposure students.

Attention-Deficit/Hyperactivity Disorder (ADHD) is a neurobiological disorder based on behavioral characteristics that can lead to "poor health outcomes, increased psychiatric morbidity, increased injury, poor peer relationships, and poor educational outcomes" (Jimenez et al., 2017, p. 356). Jimenez et al. (2017) found in their study that experiencing ACEs between ages five and nine was associated with a child being diagnosed with ADHD at age nine. They proposed this could be due to confounding symptoms of ADHD and ACEs, but either one would impact a student's ability to learn.

Gershenson and Tekin (2018) used the 2002 shootings along I-95 in Virginia as a natural experiment to study the effects of community violence on achievement. These attacks, referred to as the "Beltway Sniper" attacks, took place over three weeks and targeted random people. Ten people were killed, and three others were critically injured during this time.

Gershenson and Tekin (2018) chose to analyze academic outcomes rather than other potential consequences because they suggested these were more objective and easily observable. Other consequences of the trauma might not be as easily recognized or measurable. In their work they found school-level proficiency rates for schools within a five-mile radius of the attacks dropped between two and five percent in the area of math for third and fifth grades and also in third grade reading proficiency. Gershenson and Tekin (2018) noted this effect did reverse over time, but some social-emotional outcomes might continue even when academic scores improved.

Trauma increases the likelihood that a child will drop out of high school (Porche et al., 2011). Iachini et al. (2016) analyzed the ACE scores of ninth graders repeating the grade in three urban high schools. The participants mostly identified as African-American, and over half (54%) received free or reduced lunch. Iachini et al. (2016) found that of these students, 85% had experienced at least one ACE, and 8% had experienced five ACEs. Sixty-two percent had experienced household incarceration, and 23% had household members using substances. The authors noted that those students who had reported at least one ACE had also had a school-related behavior change whose timing aligned either at the time of the ACE or right after it.

While students who have experienced trauma may qualify for an individualized education plan (IEP) or a 504 Plan, schools do not always know if a student has experienced trauma. The student may not qualify academically for either of these but still need the support they offer through specialized instruction, accommodations, and/or behavior supports (Rossen, 2018). If schools are aware of a student's ACEs, this may help student qualify for these supports, which could positively impact their learning and grades.

Bias

Bias is an inclination for or against something and can be positive or negative (Psychology Today, n.d.). Research has shown bias exists in grading practices (Gershenson, 2020; Guskey, 2006; Hochweber et al., 2013; Link & Guskey, 2019). Grades influence the post-secondary choices students have and create a sense of self-efficacy toward content areas, and this can impact choices that last a lifetime, such as the amount of money they can earn over their lifetime.

There are two major sources of bias in grading: bias by design and bias from personal factors. Bias by design includes the choice of assessments, mismatched alignment to learning expectations, and how the teacher uses the assessment whereas bias from personal factors is related to the teacher's personal bias toward a student for any reason, such as race, gender, familiarity, or special education status (Hardré, 2018).

Teacher Perception of Student Ability. Teacher perception of how the student will perform impacts what grade the student receives (Andersen, 2018). Teachers may be more critical of students who they believe should perform at a certain level (Hardré, 2018). For example, if a student has previously scored very well on an exam, they would be less accepting of minor errors on their other pieces of work compared to a student who did not perform as well on the exam.

Green et al. (2007) found in their study that 21.9% of pre-service teachers and 34.5% of practicing teachers believed it was ethical to offer extra credit opportunities to all classes except advanced courses for high achieving students. Additionally, they found 18.4% of pre-service teachers and 23.6% of practicing teachers agreed it was okay to not give many grades of "A"

because “students’ work is rarely perfect” (Green et al., 2007, p. 1007). Both findings show evidence of bias in grading.

Hattie (2009) found that teacher estimates of achievement, or teacher belief in students, has the third highest positive impact of the 252 factors related to student achievement (as cited in Neer, 2018). It is hard to determine if this achievement is reflective of student knowledge or bias in grading. How students perceive teachers feel about them and their abilities can also impact student self-concept.

Andersen (2018) found that teachers changed academic expectations of students based on the ability level of their track placement. Teachers adjust their teaching style according to their perception of the ability level of the track, and they also increase their emphasis on acceptable classroom behaviors and discipline for tracks they view as having a lower ability. Additionally, teachers have stricter grading practices in place for those students they determine are high achievers (Brookhart et al., 2016).

Hardré (2018) suggests teachers may grade special education students based what they believe the student can do or produce rather than grading based on the performance relative to the objective criteria. This is true even for special education students who are in higher level or mainstream courses.

Race. President George W. Bush used the phrase “the soft bigotry of low expectations” in a speech given to the NAACP in 2000 as he kicked off the No Child Left Behind Act (Rubel & McCloskey, 2019, p. 115). This act aimed to close the achievement gap among various student groups with federal accountability by disaggregating student data by certain characteristics, such as race, socioeconomic status, and disability status (Ansell, 2004).

Racial bias affects student grades. Quinn (2021) found White female teachers graded work differently if they believed the student was a Black student or a White student. The study first asked participants to decide if the student had written an appropriate grade-level response and then asked them to grade the paper using a detailed rubric. The Black student received lower marks from White female teachers in both situations.

This finding is important because of the demographics of teachers. The most recent data available from the 2017-2018 school year showed that public school teachers in the United States are 79.3% non-Hispanic White and 76.5% female (Taie & Goldring, 2020). Quinn (2021) implicates race bias as having a negative impact on Black student grades, and most Black students have White teachers based on the statistics.

Conversely, Gershenson (2020) found that high grading standards could positively impact learning of Black and Hispanic students. Gershenson defined high grading standards as a practice where students receive an accurate grade reflective of their knowledge that does not show signs of inflation or inaccuracy. A teacher with high grading standards does not lower expectations. When teachers had high grading standards, it increased learning of all students, but it increased the most for Hispanic students.

Yeh (2019) posits the origin of the achievement gap is a combination of socio-cultural and socio-economic factors combined with parental factors. Practices in K-12 education exacerbate these instead of mitigating the discrepancy. Yeh (2019) notes demoralizing grading practices and grouping as instigators of the achievement gap. Non-academic factors such as behavior also contribute to this gap.

Link and Guskey (2019) stated, “When teachers interpret student behaviors through the lens of race, credit for behaviors such as being seated when the bell rings, following directions,

cooperation, and dressing appropriately may be inequitably assigned” (p. 13). The authors add this could lead to White students being rewarded with points or grades for their behavior while Black students either receive less of a reward or receive none.

Gender. In addition to race, an achievement gap has been identified between male and female students. This gap appears in classroom grades and in standardized tests. Prior research has shown that girls receive higher grades overall than boys and higher scores on reading/English Language Arts assessments, but boys typically receive higher grades in math and science (Link & Guskey, 2019; Reardon et al., 2018; Schuster et al., 2021).

Link and Guskey (2019) believe part of the explanation for the gender gap is related to non-cognitive skills. They propose that girls traditionally show better social skills and compliance with classroom behavior expectations than boys do, and this contributes to the discrepancy in grading. Classroom settings often create an environment where “boys’ less developed self-discipline skills leave them at a disadvantage” (Link & Guskey, 2019, p. 14). If behavior is a factor in grading, this can have an impact on grades; however, teacher perception of acceptable behaviors can also unconsciously impact grading.

The format of assessments can also create a gender bias. Reardon et al (2018) conducted a nationwide study where the standardized state accountability tests of roughly eight million fourth and eighth grade students were analyzed to determine if a gender gap existed. They found boys outperform females on state accountability tests where the questions are more multiple choice rather than constructed response or essay format. They found this to be true in both math and reading. Although the gap was smaller in math, it was still statistically significant.

The authors concluded the differences in student performance between boys and girls was “large enough to have meaningful consequences for the students” (Reardon et al., 2018, p. 291).

The ACT is currently entirely multiple choice (ACT, 2023). Tennessee end-of-course assessments are all multiple choice or multiple select except for the English I and English II assessments. These both contain an additional subpart that requires students to respond to a writing prompt (Tennessee Department of Education, n.d.-c). It could be deduced boys outperform girls on these assessments based on the formatting of the assessment.

Gender bias is more pronounced when the teacher is the gender typically associated with that content area. Female teachers who taught math, which is stereotypically masculine, and male teachers who taught languages, which is stereotypically female, were found to be less likely to show gender bias in their grading (Doornkamp, 2022). The researchers theorized this information contradicted other researchers who believed gender stereotypes are less amenable but also stated the effect size of their study was fairly small.

Familiarity of Student. In communities that are tighter knit, it is not uncommon for teachers to have a relationship with students and/or their families outside of their classroom. They may have taught multiple family members, and they are often more embedded in the community (Hardré, 2018). This proximity can lead to bias in grading. Knowing students or their families well can evoke a positive or negative bias in grading (Hardré, 2018). If a family has a history in the community of being local leaders or if the teacher knows the students' parents as a professional colleague, they may tend to grade with more leniency compared to a student with older siblings with whom the teachers previously had a negative experience (Hardré, 2018; Malouff, 2008).

Local Grading Policy

School boards act as an “elected public body with authority to set direction for the school system” (Tennessee School Board Association, 2016, p. 4). State law requires these boards to

develop and publish written policies that are reviewed every two years. Tennessee School Board Association policies are meant to provide clarity to the public, save time in making decisions, and help the board act appropriately and legally.

The published grading policies of seven school systems in upper east Tennessee were reviewed. The researcher specifically examined grades 9-12 policies. All seven systems use an A, B, C, D, F system aligned to a range of percentages (see Table 2). Additionally, all seven systems have a policy for points added to Honors, Dual Credit, Dual Enrollment, Capstone, Advanced Placement, Cambridge International, College Level Exam Program, and/or International Baccalaureate course: Honors receive 3 additional points; Dual Credit, Dual Enrollment, and Capstone receive 4 additional points; and, Advanced Placement, Cambridge International, College Level Exam Program, and International Baccalaureate courses receive 5 additional points. These points are added to calculate the semester average.

One system states that students must take the Advanced Placement exam to receive five points; otherwise, they will receive three points, which is the amount earned in an Honors course. One of the seven systems includes the requirements for what must be included in Honors courses. There are no requirements listed for the remaining six districts.

Two of the seven systems also weight the quality points earned: Honors courses receive 0.5; Dual Credit, Dual Enrollment, and Capstone receive 0.75; and, Advanced Placement, Cambridge International, College Level Exam Program, and International Baccalaureate courses receive a 1.0. Only one system allows for + or – to be added to the alpha grade, such as A+ or an A-. The other six systems either state plus/minus evaluations are not to be used with grades or do not include it in the adopted grading scale.

Two of the seven districts state how grades are to be determined:

Grades given at the end of each nine (9) weeks period will be determined from daily work, homework, written assignments and tests. The teacher will weigh the value of grades given for various assignments and tests within the applicable period in computing the grade. This procedure will enable the teacher to allow for individual student differences in the grading process. Any assignments and tests required of a student must be considered in the computation of his grade (Greene County Board of Education, 2022; Johnson City Board of Education, 2022).

While teachers can determine the appropriate weighting of various assignments and tests, the policy states any required assignments and tests must be included in the final calculation of the grade. The other systems do not include language stating requiring specific categories (daily assignments, tests, etc.) must be included in the grades. One system includes a requirement for semester exams to be calculated at 15% of the final semester grades.

All seven systems explicitly state conduct grades are based on behavior and will not be reflected in scholastic grades; however, two systems provide scales to communicate behavior to stakeholders. Two of the seven systems state attendance records will be used to determine the awarding of grades, passing a course, or promotion/retention. Four of the seven systems do not include any language referencing attendance, and one system explicitly states attendance is not to be included in calculating grades. Table 3 shows a summary of the policies.

Table 3*District Grading Policy Analysis*

Characteristic	Policy 1	Policy 2	Policy 3	Policy 4	Policy 5	Policy 6	Policy 7
Letter grades aligned to percentage ranges	X	X	X	X	X	X	X
Points added for Honors, Dual Credit, Dual Enrollment, Capstone, Cambridge International, College Level Exam Program, Advanced Placement, or International Baccalaureate courses	X	X	X	X	X	X	X
Advanced Placement exam required for additional points					X		
Framework Standards Required for Honors courses					X		
Weighted quality points for Honors, Dual Credit, Dual Enrollment, Capstone, Cambridge International, College Level Exam Program, Advanced Placement, or International Baccalaureate courses				X	X		
Additional weighted quality points for Honors and Advanced Placement courses based on nine weeks grades					X		
Categories required for grade calculations		X		X			
Semester exam required					X		

Characteristic	Policy 1	Policy 2	Policy 3	Policy 4	Policy 5	Policy 6	Policy 7
Conduct scale included in policy			X		X		
Conduct Included in Grade							
Attendance included in grade				X	X		
Attendance not included in grade			X				
Attendance in grade not addressed in policy	X	X				X	X

Theoretical Framework

There are many theories to explain human behavior. Social Cognitive Theory postulates that people are “neither driven by inner forces nor automatically shaped and controlled by external stimuli” (Bandura, 1986, p. 18). This theory originated from the work of Lev Vygotsky who studied human development.

Vygotsky theorized this development took place on two planes: the intermental and intramental (Eun, 2019). On the intermental plane, people engage in a social interaction, and the skills and knowledge gained from the first plane become internalized on the second plane, the intramental plane (Eun, 2019). This process requires a more knowledgeable or capable person(s) to lead the less competent person(s) during the social interaction in the first plane so that the less competent person(s) can internalize the skill or knowledge to use for their own cognitive processes. This development became known as the General Genetic Law of Cultural Development, and it explained that this applies to all human forms of behavior.

Bandura’s Social Learning Theory, which is the origin of Social Cognitive Theory, complements Vygotsky’s General Genetic Law of Cultural Development. It is centered on the notion that humans learn by observing the actions of others and the positive or negative

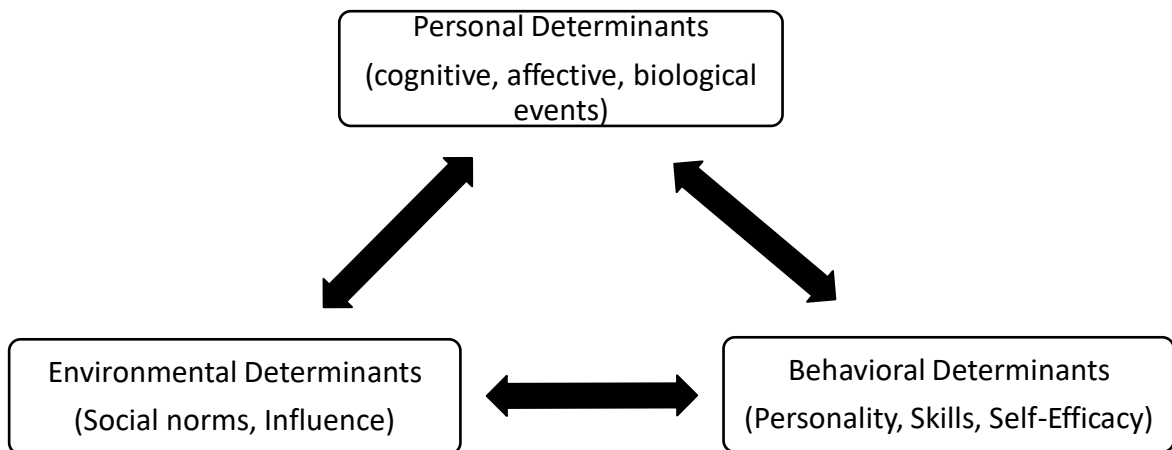
consequences that result from those actions (Lyons & Berge, 2012). Bandura (1977) theorized that people learn through everyday events what behaviors are successful and which ones are not. Those that are not successful are rejected.

According to Bandura’s Social Learning Theory (1977), humans learn most of our behaviors by observing others modeling these behaviors. The behavior, how it is performed, and the response or outcome it elicits are stored cognitively for future use. This contrasts with other theories that tried to explain human behavior because of internal impulses (Bandura, 1977). Bandura (1977) states, “Rather, psychological functioning is explained in terms of a continuous reciprocal interaction of personal and environmental determinants. Within this approach, symbolic, vicarious, and self-regulatory processes assume a prominent role” (pp. 11-12).

In his Social Cognitive Theory, Bandura analyzed human behavior as an interaction among three components, referred to as triadic reciprocity (Bandura, 1986). In triadic reciprocity, personal determinants (cognitive and emotional), behavioral determinants, and environmental determinants fluidly influence each other during social interactions (Ferrari et al., 2010). Figure 1 shows this interaction among the three determinants.

Figure 1

Triadic Reciprocity



Note. Adapted from *Social Foundations of Thought & Action: A Social Cognitive Theory* by Albert Bandura, 1986.

When two or more people engage in a social interaction, each party involved will experience response consequences, both positive and negative, because of this exchange. Response consequences serve three functions: They are sources of information, act as motivators, and strengthen responses. As a source of information, response consequences affect cognition and how people assign thoughts, which are sometimes subconscious, to an action. The motivational function of response consequences causes people to act on prospective events. Prior actions generate a classification system where those behaviors are ranked by their effects. They can generate benefits, have no impact, or “avert future trouble” (Bandura, 1977, p. 18).

Bandura (1977) did not incorporate the third function of strengthening but does advance regulation as a modifier of behaviors. Additionally, Eun (2019) notes the impact of the environment affects people differently based on where they are developmentally at that phase. Teachers who are at the beginning of their careers may be impacted differently than teachers who are considered to veteran or expert teachers.

Vygotsky’s General Genetic Law of Cultural Development, and Bandura’s Social Learning Theory and Social Cognitive Theory can all be used as a lens to examine human behavior and explain the decisions made by them.

Summary

This chapter is an overview of the existing research related to grades grading practices. The literature chosen provides a history of the foundations of the educational system from the earliest days to post-2000, explores the purpose of grades and grading practices, explains the common components of grades, and describes the significance of measurement. Several different

grading systems are explored in addition to grading scales and grading practices. Within grading practices, the literature delves into factors, such as mindset, bias, and trauma, that can affect grades and grading practices. The theoretical frameworks for this study, Social Learning Theory and Social Cognitive Theory, are also addressed in this chapter. Chapter 3 describes the methodology for this study. Chapter 4 details the data analysis and its findings. Chapter 5 summarizes the findings, includes implications of the research, and recommendations for future research.

Chapter 3. Methodology

The purpose of the phenomenological study was to understand the educator experience of choosing grading practices for math and English courses East Tennessee high schools.

Research Questions

To understand the educator experience, the researcher developed a singular research question to guide the study: What are the lived experiences of high school math and English teachers when choosing grading practices in their classrooms?

Supporting Research Questions

Several research questions to answer the essential research question were developed.

RQ1: How do teachers perceive the use of grades in secondary math and English classrooms?

RQ2: What experiences are associated with choosing a classroom grading practice for secondary math and English classrooms?

RQ3: What experiences have teachers had with professional development centered on grading practices and policy for the classroom?

RQ4: What external influences affect teacher choice in classroom grading practices?

Qualitative Research Design

This research study is a phenomenological study focused on deepening the understanding of teacher experiences when selecting a classroom grading practice. Phenomenology is a qualitative research tradition that seeks to compress multiple experiences of those who have lived it (Merriam, 1995) into a phenomenon into a “universal essence” (Creswell et al., 2007, p. 252).

Phenomenology, commonly used in sociology, nursing, education, and health sciences, is strongly influenced by the German mathematician Edmund Husserl (Creswell & Creswell, 2018;

Creswell et al., 2007). Researchers who use phenomenology are interested in the behaviors of those who lived the phenomenon from their perspectives (Tenny et al., 2021). Within this tradition, there is hermeneutical phenomenology, which focuses on the interpretation of the researcher, and transcendental phenomenology that relies on the experiences of the participants (Creswell & Poth, 2018). This study is rooted in transcendental phenomenology.

To understand the perspective of the person who lived the phenomenon, the researcher must remove themselves from the process. Epoch is the first step in this process. The researcher reflects to find any personal bias and gain clarity about any personal feelings (Patton, 2002). This contributes to the rigor of the study. Next, the researcher examines the phenomenon separately from any external influences. These first steps are to provide a fresh view of the phenomenon (Creswell et al., 2007; Creswell & Poth, 2018).

Data were gathered through interviews and analyzed for themes. The data is horizontalized, which is the process of giving each piece of data equal weight, identifying significant statements, and finding common, meaningful clusters among the data (Creswell & Poth, 2018; Patton, 2002). The researcher then develops themes to write the description of what the participants experienced and reports the essence of the experience (Creswell et al., 2007; Creswell & Poth, 2018).

This research was designed to understand the experiences of educators when selecting a classroom grading practice. This research paradigm posits that reality is constructed by individuals and their “experiences, interactions, and backgrounds” (Tenny et al., 2021, para. 11). This paradigm rests in ontological relativity that theorizes two people can live in the same empirical world and have completely different views of this world (Patton, 2002; Tenny et al.,

2021). This model allows the researcher to ask ‘how’ and why’ questions around the phenomenon (Darlaston-Jones, 2007).

Site Selection

This research study was conducted at rural high school in East Tennessee. The school serves grades 9 – 12 and has an enrollment of 932 students. The graduation rate is 97%. The researcher chose to conduct the study within one school to reduce district-specific factors such as board policies from affecting the data. School systems often have a district focus for professional development and focusing on one school allows the data to be analyzed through a specific lens without the influence of different district goals. The school’s pseudonym for the purpose of reporting the findings is Addie LaRue High School.

Participants

Leedy and Ormrod (2019) stated “the quality of the data obtained can be only as good as the quality of the sample(s) used” (p. 270). This study used purposive sampling because it was necessary that the participants had experienced the phenomenon to be able to provide thick, rich data (Creswell & Poth, 2018). Because this study focused on educators with a shared experience of choosing classroom grading practices, only persons who met specific criteria were included in the sample. Participants had to meet the following criteria: must be a certified educator in the state of Tennessee with a minimum of three years of experience and must teach high school English or math. These were chosen as the content areas because more of these courses are required for graduation in Tennessee, and the course work follows a progression with pre-requisites. The English department is composed of nine teachers, and there are 11 teachers in the math department.

Qualitative research generally has a very small sample size so that the participants can provide an in-depth understanding of the phenomenon (Merriam, 1995; Patton, 2002). Creswell and Creswell (2018) reviewed different qualitative designs and found that the sample size of phenomenological studies ranges from three to ten. The sample for this study included three high school math or English teachers with at least five years teaching experience. Although the sample size for this study was small, it did achieve saturation. Time constraints and scheduling conflicts did not allow as educators to participate as was initially anticipated.

Educator 1 has been teaching for 25 years, and their highest level of education is an Educational Specialist degree. They also have National Board Certification from the National Board for Professional Teaching Standards. Educator 2 has been teaching for 12 years, and their highest level of education is a master's degree. Educator 3 has been teaching for 26 years, and their highest level of education is an educational specialist degree. Educators 1 and 3 followed a traditional path to the classroom, whereas Educator 2 followed a non-traditional path where they returned for a master's degree to teach.

Data Collection Strategies

Qualitative data should “capture and communicate someone else’s experience of the world in his or her own words” (Patton, 2002, p. 47). To capture an experience of another, it is essential to develop interview questions that seek to support the goal of the study and allow participants to share the information necessary to gain understanding of their lived experiences.

Interviews were conducted with the participants so they could express, in their own words, their personal experiences, and open-ended questions were designed to elicit data that conveyed the essence of their experiences. The researcher used memos during the research

process to reflect on and use as a tool to determine how their individual experiences may affect the interpretation of the data.

Interviews

Interviews are employed to gather data when the researcher wants to learn things that cannot be observed (Patton, 2002). The researcher should create a safe environment where the participant feels comfortable sharing their experience. They should also design questions that allow the participant to provide prolific data and listen intently to the responses (Qu & Dumay, 2011). The ability of the interviewer to do these things greatly impacts the quality and depth of the data.

Interviews were conducted with each participant individually. A conversational format was used to allow for more open communication (Creswell & Poth, 2018). Because no other interviewers were involved, the researcher was able to ask probing questions appropriate to the research when applicable. Participants were also given the opportunity to share any additional information at the conclusion of the interview questions. The interviews took place on Zoom, a videoconferencing software. The participants chose their location, and each one chose their classroom. They were recorded, and these audio files were used for transcription. The audio files and transcriptions were stored electronically and were password protected both on the device and in the cloud account of the researcher.

Data Analysis Strategies

Data analysis is not an independent step but is part of a progression used by qualitative researchers (Creswell & Poth, 2018). Data collection and writing are often happening simultaneously as the researcher seeks saturation. Creswell and Poth (2018) describe the data analysis as a spiral where the “researcher engages in the process of moving in analytic circles

rather than using a fixed linear approach” (p. 185). This spiral includes managing and organizing the data, reading and memoing emergent ideas, coding the data and developing themes, relating the themes to the framework, and reporting the data.

The researcher utilized a social constructivism framework using Bandura’s Social Learning Theory and Social Cognitive Theory and pursued complex views rather than narrow meanings of the phenomenon (Creswell & Poth, 2018). The research is, therefore, dependent upon the views of the participants.

The researcher recorded the interviews to assist with transcription. All interviews were transcribed from these recordings for the purpose of coding the data. Open and axial coding was used. In open coding, the data is broken down into discrete segments and common codes or themes are identified (Scott & Medaugh, 2017). The next step in this process is axial coding, where the researcher makes connections among these segments or codes. These processes developed themes by aggregating the codes for the researcher to present in the findings (Creswell & Creswell, 2018; Creswell & Poth, 2018).

Assessment of Quality and Rigor

Trustworthiness

Validity and reliability are addressed in qualitative research as trustworthiness (Amankwaa, 2016). There are four criteria to establish trustworthiness: credibility, transferability, dependability, and confirmability. Credibility is the amount of confidence the reader has that the data represents the truth, and transferability is how applicable the research is in other contexts (Amankwaa, 2016).

Dependability shows the findings are consistent and repeatable, and confirmability describes how neutral the researcher is regarding “bias, motivation, or interest” (Amankwaa,

2016, p. 121). To ensure trustworthiness of this study, the researcher triangulated the data, provided thick, rich descriptions of the data, and used reflexivity to filter any bias that might be present in the study.

Thick Description

Thick, rich descriptions were used to provide the reader with a deeper insight and to give “the discussion an element of shared experiences” (Creswell & Creswell, 2018, p. 200). Details, such as description of physical, movement, and activity details, support this by transporting the reader into a realistic setting (Creswell & Creswell, 2018; Creswell & Poth, 2018). Additionally, moving “from general ideas to the narrow, interconnecting the details, using strong action verbs, and quotes” provides a thick, rich context that allows the reader to transfer this study to others (Creswell & Poth, 2018, p. 263). The researcher included quotes from the participants to support thick descriptions of the research.

Triangulation

Triangulation of the data examines additional evidence to build justification for the themes found by the researcher and helps the researcher understand any inconsistencies that may be present in the findings (Creswell & Creswell, 2018; Patton, 2002). The participants were asked to review the findings to add context or offer an alternative interpretation of the data to support its triangulation.

Reflexivity

In qualitative research, the researcher interprets the data in relationship to their own experiences that they bring into the study, and all writing is situated within these experiences (Creswell & Poth, 2018). The researcher should convey to the reader how these experiences have shaped their interpretation of the data through reflexivity, which is sharing past experiences with

the phenomena and being explicit about how these experiences have influenced the interpretation (Creswell & Creswell, 2018; Creswell & Poth, 2018). To do this, the researcher must be self-aware. This is “to undertake an ongoing examining of *what I know* and *how I know it*” (Patton, 2002, p. 64).

Ethical Considerations

The researcher received approval for permission to conduct this research from the Institutional Review Board (IRB) of East Tennessee State University. As part of the IRB approval process, the researcher submitted and received approval for the following forms: permission from the school district to conduct research within the district; recruitment email; and consent from the participants. Participant consent was obtained and documented before conducting interviews.

Summary

The methodology described in this chapter is related to the essential question of this study: What are the lived experiences of high school math and English teachers when choosing grading practices? This qualitative study involving three high school math or English teachers used interview data based on four research questions developed to support the overarching focus of the study. Open and axial coding was used to analyze the data, and social learning theory was applied to the emerging themes that developed during coding. The assessment and quality and rigor are included in this chapter. Chapter 4 presents the finding of this study in relationship to the research questions and emergent themes. Chapter 5 provides interpretations, conclusions, and recommendations about the data for practice and future research.

Chapter 4. Findings

Grades and grading practices are multi-faceted and meant to serve many purposes (Olsen & Buchanan, 2019): They communicate an objective measure of a student's quality of work or mastery of learning; they judge a student's effort of work, willingness to participate, or ability to follow directions; and they compare students against a standard and to each other. Grades are used to inspire and punish. They can do all these things in multiple combinations, but they are not able to do them all at the same time.

Qualitative interviews were conducted with three high school educators. Two of the educators currently teach math, and one educator teaches English. Two of the educators were female and one was male. One teacher has taught for 12 years, while the other two have taught for 25 and 26 years respectively. All three participants have taught in more than one school system, and one has taught in more than one state. Two educators have obtained their educational specialist degrees, and one has their master's degree. One educator has also obtained national board certification in addition to the educational specialist degree.

The data is arranged by the research questions and the themes that emerged from each. Three additional themes emerged from the body of data and those findings are shared after Research Question Four as Additional Themes. A pseudonym has been used for the name of the school, and participants are referred to by a number (e.g., Educator 1).

Purpose of the Study

The purpose of the phenomenological study was to understand the educator experience of choosing grading practices for high school math and English courses.

Research Questions

To understand the educator experience, the researcher developed a singular, overarching research question to guide the study: What are the lived experiences of high school math and English teachers when choosing grading practices in their classrooms?

To guide this study, four research supporting research questions were developed to answer the essential research question.

RQ1: How do teachers perceive the use of grades in secondary math and English classrooms?

RQ2: What experiences are associated with choosing a classroom grading practice for secondary math and English classrooms?

RQ3: What experiences have teachers had with professional development centered on grading practices and policy for the classroom?

RQ4: What external influences affect teacher choice in classroom grading practices?

Data Collection and Data Analysis

Open-ended interview questions were designed by the researcher so participants could share their experiences in their own words. Participants were able to share in-depth responses to open-ended questions that may be atypical or unfamiliar to the researcher. Pre-determined responses would not elicit the same rich, thick data as compared to open-ended questions.

The educators were interviewed using video conferencing software, and the researcher used these recordings to transcribe the data. Open and axial coding was then used to analyze the data for emerging themes.

Framework

The framework used for this study was Bandura's Social Learning Theory. This framework helps understand human behaviors and explains that behaviors, such as grading

practices, are learned through social interactions and observations of others (McLeod, 2016). When a behavior elicits a positive response, it is stored as a behavior to continue; however, a behavior with a negative response is removed as a potential future response (Lyons & Berge, 2012).

Bandura (1986) theorized three determinants (behavior, cognitive, environment) individually influenced the effects of the others. He refers to this as triadic determinism. The influences of the determinants are not fixed; rather, they are consistently fluid and only dependent on the circumstances and the individuals involved in a particular activity (Bandura, 1986).

At any given interaction, one determinant may emerge as dominant. Bandura (1986) states, “If people are dropped into deep water, they will all promptly swim” (p. 24). If the environmental situations are weaker, personal preferences may more strongly impact the actions rather than the environment or cognition, such as choosing a book from the library. Cognition, or how a person thinks about something, can determine actions depending on how strong those beliefs are. If these beliefs are strong enough, the social consequences are not enough to change those beliefs (Bandura, 1986).

This framework was chosen because grades and grading practices involve a person’s beliefs about grades and grading practices, the school environment in which they learned about grades and grading practices and the school environment in which they now teach, and the positive and negative behaviors that accompany grades and grading practices.

Findings

Participants were asked several interview questions to support the research questions. Probing questions were used when appropriate during the interviews to encourage thick and rich

descriptions of the phenomena of choosing a grading practice for their classrooms. Participants were given time at the end of the interviews to share any additional information they had not had the opportunity to share. This open-ended opportunity provided additional data that would not have emerged otherwise and added to the thick and rich descriptions in the findings.

The following themes emerged from the data analysis:

1. Communication among stakeholders (student, parent, teacher, school)
2. Student Effort
3. Value of learning
4. How to grade and implement grading practices (for preservice and practicing educators)
5. External pressures
6. Grades as motivators
7. Teacher expectations
8. Reliability and validity in grading practices

These emergent themes can be categorized into three categories: personal affect, teacher affect, and environment affect. Some themes emerged across the body of data and others were isolated to a question but appeared among all participants. These categories allow the data to be considered by source of affect. Personal affect is determined by the student; teacher affect is impacted by the actions and beliefs of the teacher; and the environmental affect is influenced by factors that are a combination of student, school, home, and undefined external factors. Table 4 presents the themes by affect category.

Table 4

Summary of Emergent Theme Categories

Personal Affect	Teacher Affect	Environmental Affect
Grades as motivators	How to grade and implement grading practices	Communication among stakeholders
Student effort		External pressures
External pressures	Value of learning	
Teacher Expectations	Teacher Expectations	
	Reliability and validity in grading practices	

Note. Some themes were applicable to more than one category and were listed accordingly.

Research Question One

The first research question (RQ1) for this study asked how do teachers perceive the use of grades in secondary math and English classrooms? The themes that emerged from this question were communication among stakeholders, student effort, and value of learning.

Communication Among Stakeholders

Communication among stakeholders emerged as a theme. This communication was represented by conveying to the students how they stand in relationship to the course expectations and as feedback to the teacher reflecting the effectiveness of their teaching.

Educator 1 said communication was on-going in their class in the form of feedback, but they also had to assign a grade because that was the expectation. They stated, “So, my goal by the time I get an essay, I hope that everybody’s at an A because we did all the work upfront and all the conferencing upfront.”

Educator 2 shared, “The purpose of grades is to let a student know where they are in relation to the expectations of the course, and, so, uh, that expectation should be grade level.”

Educator 3 said, “I guess it’s the figuring out how the kids are doing. How much are they learning as you...as you teach them.”

Student Effort

The educators were then asked to describe what should be included in a grade, specifically a marking period grade or an end-of course grade. The theme of student effort emerged from the data.

Educator 1 stated:

What I include, obviously, I mean, I, their work that they did, but I also include – especially if it’s a kid that is borderline – effort and attitude. If a kid’s on the border, and they were like, so sweet or they worked so hard, or they might even have failed, I might pass them, but if a kid’s like a real jerk the whole time and they’ve got, like, a B but they could have an A...(shrugs shoulders) they might sit on that B. So, I do think attitude and effort in general pleasantness can help in my opinion. Yeah, I think it does....Yeah.

Educator 2 shared they felt their school’s system was like what they had experienced in other schools and was “not bad”. Their department weights grades at 60% for tests, 20% for quizzes, and 20% for homework; however, there are no expectations for how many of those are given in a marking period except for a department-wide mid-term and the EOC required by the state in select courses.

Homework is graded for effort (completion) in their department and not for accuracy. Regarding this, they stated:

Probably grading behavior more than anything. That 20% of homework and classwork...grading the fact that they're keeping up with their work, that they're, you know...I grade on completion, as do most of the teachers around here, grade homework on completion versus accuracy. (...) I go over in class, so they do get that feedback, but it's a pretty...It's a pretty safe way to practice. You know, for them. You know, the last thing I want to do is, like, stress the child out over something that they just learned, too, kind of thing.

Educator 3 noted there were not a lot of options to be included in grades, and they believed homework should be graded. They suggested using the method they had used in high school where papers were exchanged between students and they each graded another student's paper. They felt this was efficient and allowed students to ask questions about their work in class.

The current practice for their department is to grade homework based on completion. They explained that while students might be putting in effort to complete the work, not all students were putting in genuine effort to learn the material.

They stated, "In a perfect world, the student has to take some ownership in wanting to learn, so they're willing to ask questions. I can answer questions till I'm blue in the face and still not answer yours." They shared with the researcher that this belief may be due to subject area and that they were "old school".

Value of Learning

The value of learning also emerged as a theme from this research question. This theme was shared through different lenses: Opportunities to re-do work so students learn the material; a

student deciding if there was value in his homework; and explaining where the learning takes place in one educator's classroom.

Educator 1 said, "So, really, I...and then, if they make a bad grade on a test, we go back and re-do because I really want them to learn, but I have to give a grade." Educator 2 shared a student perspective:

But, you know, I had somebody at the end of last year who was, like, "I feel pretty confident that I'm going to get a 5 on the AP exam, so I'm not gonna do all of the homework" and that kind of thing. So, I mean, he understood what impact that had on his grade, and you know it's one of those things where they're edging toward being a college student that would be making those...those decisions as well. So, you know, it is what it is.

Educator 3 also spoke to the learning in the form of homework. They said, "Taking your homework a little more serious because, in reality, that's where the learning takes place." This educator expressed frustration because they felt their department policy of grading homework on completion/effort was detrimental to learning.

Research Question Two

The second research question (RQ2) for this study asked what experiences are associated with choosing a classroom grading practice for secondary math and English classrooms? The theme of how to grade and implement grading practices emerged from this. All three participants shared that their knowledge of grading and eventual grading practices came from either a mentor teacher or colleagues. One remarked they did not know it was necessary to learn about grades and grading, and two conveyed that the only preservice experience they had was restricted to creating an assessment with the emphasis being on writing test questions.

How to Grade and Implement Grading Practices

When asked what preservice experiences they had had, Educator 1 answered it was very little: “I bet we learned about rubrics. I’m going to say that’s all. I’ve...in my master’s degree, I learned about how to, um, properly word a question on a test (...) but other than that, literally zero. Rubrics were the *thing* [emphasis added].”

Educator 2 followed a non-traditional path to the classroom, and they did not have any preservice experiences with grades or grading practices except for learning how to create an assessment until their student placement in the public school. They shared that at their first school as a practicum student the school put emphasis on passing the students, and the school had a one through four grading system. “They had to, like, the kid only had to get a 25% to get a D because of how they did grading. It was a mess.”

Educator 2 completed their student teaching at a different school, and they expressed they had much more autonomy there. “And so I had an awesome mentor teacher there, and so her style of grading kind of became my style of grading.(...) And then, you know, I’ve seen how people grade here and people grade in other cases and kind of taken from that.”

Educator 3 expressed that they thought grades and grading practices were just understood by educators: “I just assumed that’s kind of an intuitive things that that’s what you’re going to do.” They further explained each person added their own influence to their practice: “You know, one teacher may write down a test grade twice in the grade book and another teacher would write it down once and put a percentage with it.” Once they entered the classroom, they said grading practices were developed based on asking other teachers what they were doing. They based their grading practices on what they had experienced as a student and what “older” teachers were

doing when they began teaching. They summarized their experience by saying, “But I mean, this, I never really thought about it needed to be told how to do it.

Research Question Three

The third research question (RQ3) for this study asked what experiences teachers had with professional development centered on grading practices and policy for the classroom? The theme of how to grade and implement grading practices also emerged from this question. Each participant shared they either had not had opportunities for professional development related to grades grading practices or they had minimal engagement with them.

How to Grade and Implement Grading Practices

Educator 1 responded that her only professional development related to grades or grading practices was standards-based grading. They continued, “I’m just not interested in it.” The interview was interrupted for about 30 seconds by another educator in Educator 1’s building, and she asked for the question to be repeated. The researcher responded by reminding them that we were discussing learning new things related to grades and grading practices. She responded, “Right. I mean, I’ve got my master’s. I’ve got my Ed.S. That was never really mentioned except that one thing in my master’s that I had to create a good test.” While this answer referenced preservice experience, it did point out they had not had professional development for grades or grading practices other than the exposure to standards-based grading.

Educator 2, similar to Educator 1, had participated in standards-based grading professional development, but it was not content specific. They stated:

I’m not sure where they landed on that cause it’s, you know, that’s a whole other beast, and, you know, it’s not that I’m against standards-based grading. It’s just, you know, it’s

just a big shift. (...) It seems like a whole school needs to do or the whole school doesn't do it.

Educator 3 did not express knowledge of any professional development for grades or grading practices: "No, I've never thought about it, to be honest. It's...I don't...I don't think I've ever seen anything along those lines."

All educators teach at the same high school but in different departments. Standards-based grading was a goal of a previous administrator, and school and district leaders had visited other school systems who had successfully implemented standards-based grading. The administration has since changed with none of the original three administrators or director of schools still in those positions.

Research Question Four

The fourth research question (RQ4) asked what external influences affect teacher choice in classroom grading practices? The researcher asked educators about their school's culture around grading and what external factors affected grades and grading practices. The emerging theme from this research question was external pressures. All three educators spoke repeatedly to the external pressures on students, teachers, and schools.

External pressures

Educators were asked to describe the culture of grades in their school. Educator 1 articulated that students just want a grade and are not concerned with the learning. They spoke to honors students specifically in their response: "Hmph, well, again, if you're an honors kid, all they want is the A. It's all the parents care about. It's all the kids care about. By any means possible. Not learning."

Educator 1 had spoken to parents several times in previous questions. When asked about external factors, she vigorously shook her head in the affirmative and said, “Parents. And I try not to make them mad.” While this educator’s initial response spoke to the theme of external pressures, they continued by explaining the relationship between external pressures and external factors. They added that honors parents placed pressure on teachers and students for high grades while other parents did not.

People that want to go to college. They will cheat and lie to get in. Whereas these people are like whatever, but they learn it. (...) That’s why when I came to Addie LaRue , I said I’ve got to get out of honors. I’ve got to take a break. I did it for 15 years, and I just could not pander to anybody anymore.

Educator 2 communicated they felt parental pressure also impacted culture. They remarked they felt this had been increasing throughout their 12-year career, and students and parents were much more concerned with the grade point average than they have been in previous years. They shared a personal example of this where a student was moved into their class for the spring semester after having a different teacher in the fall semester. The student and parents had complained about his grade, so he moved down an ability level to their course. Their course was taught online as this was during 2020-2021 when COVID-19 impacted schools’ learning delivery. They theorized that he wanted to move to online because he thought it would be easier.

They recounted:

And so, at one point, you know, this child was not in any honors classes, and that’s not a bad thing, but, I mean, it wasn’t like he was, you know, very, very high achieving in general, and at one point, the dad says to me, “You’ve ruined it all. You all have ruined any chance of a scholarship for him.” And I’m like...and, like, it’s January of his

freshman year. Like, you know, I mean, there's just like, there gonna be number one harder stuff than this, and number two, like, he's not, like, he's not in the upper, upper echelon of everything, so, like, it's gonna be ok. Like, it's gonna be ok, but trying to get that across is sometimes difficult.

Educator 2 also teaches honors courses and voiced concern with these students. "...I think there is a point where some parents and some students are just so grade-motivated because they are so deathly afraid of what might happen if their GPA isn't what's in their head." They went on to share that some parents had expressed outrage when a student had received a zero on a test for cheating even though it is written into the policy and handbook that students and parents are supposed to read and sign annually.

When asked how students felt about the culture of grades, Educator 3 expressed students did not think about grades and only worried if they feared backlash from a parent, but added most parents did not care, either. Educator 3 further shared their personal experience from a parent perspective with the pressure of grades.

They shared an experience with their child who had taken more challenging courses in high school and had made a B in a rigorous course. They theorized this same child would have made an A at another school but had made a B here because the rigor and expectation level were extremely high. They said this caused students with a slightly higher GPA to have a better chance at scholarships than their child even though their child might have more knowledge. They continued that this was additional pressure on their child because the child felt she would be competing at a disadvantage against other students from different schools for scholarship money or admission to schools.

Educator 3 expressed the high-stakes nature of grades on a professional and personal level. Speaking on a professional level, they expressed students felt pressure to receive good grades. “Everybody wants a good grade, and there’s so much pressure to give me a good grade. Nobody wants the bad grade.” This pressure had negatively impacted students in their view.

They said:

In the spirit of No Child Left Behind, in the spirit of everybody makes it, great graduation rate being the number one most important thing, the pressure on the school system to have good test scores, the pressure to graduate at a certain percentage... We’ve put so many scaffolds underneath them, they don’t know how to fail. They, they don’t even have to hold on. The scaffolding holds them up there. They don’t...they never get to experience what happens if you don’t do what you’re supposed to do, and it’s never your fault if you don’t do it.

Additional Themes

The themes of grades as motivators, teacher expectations, and reliability and validity in grading practices also emerged. These themes were not isolated to a singular research question, and they appeared as frequently as the other themes once the data was analyzed and were included for this reason.

Grades as Motivators

This theme emerged from two different perspectives: Two educators shared their experience with grades as motivators from the lens of high-achieving students being motivated by high grades, but one educator shared their experiences with students being motivated by lower grades in hopes of an easier course.

Educator 1 spoke to two different types of students. They acknowledged the positive motivation grades for those they described as good students, but they also expressed good grades were not a motivator for other students. Educator 1 stated, “Honestly, it is, if you want my real opinion, it is a motivator for your good students, and somewhat of a punishment of the kids who don’t really care because we do everything in class.” They went on to say, “I mean, it’s not really what I’m supposed to say, but that’s what I really think. But that’s what I *really* [emphasis added] think.”

Educator 2 revealed in their interview that they were a very grade-driven student when they were in high school and how that affected them as an educator. They remarked:

As a student, as like a high school student, I was very much grade-driven. Like, I was like the kid who was grade-driven. I wish I could have figured out that that wasn’t the...that wasn’t what I needed to be, but that’s what it was, you know.

They later connected their personal experience as a grade-driven student to their current Advanced Placement (AP) students.

And, to be honest, sometimes I will look at, like, my AP kids and see have they had a B before. Because if they had a B before, they were much more cooperative, much more into the learning generally...generally, than a kid who’s never had one.

Educator 3 spoke to grades as motivators from a different viewpoint. Their perspective was that students were not motivated in a positive way by grades, but the fear of a grade that was not acceptable to their parents was the motivator. “Today’s culture, they could care less. They don’t care about the grade until there might be some backlash from a parent.”

Educator 3 expressed the culture centered on homework being about effort and not learning. They reflected on the effort-based practice adopted by their department:

And as long as the effort being is being put into it, then it's great. If there's not a lot of effort being put into it and we're just turning stuff in, then, which is...human nature has a tendency to do when you're not getting graded on it every day.

In addition, Educator 3 conveyed frustration with different groups of kids related to effort:

You still have, lack of phrasing, the better kids do what they're supposed to, and the other ones don't. And, but, that's ...the problem is we've...My opinion is the bad apple spoils the whole a lot more than a good apple will make the apples good. You know what I'm saying? That's that. When you start giving freebies or easy to the lower level, the level right above that wants the same thing and it, that soreness, for lack of a better phrase, it spreads like rot does, and a bunch of apples.

Educator 3 later added that some students intentionally tried to do poorly so they could be placed in an online credit recovery program that takes place during the day in the school building. Students go to classes where regular classes are taking place and work on their computers independently in the back of the room. Once students reach a 60% (D), they receive credit for the course toward graduation and are moved to a different class or continue with another credit recovery course in the same setting. They stated, "Getting through it to...all they're...all they're interested, it's getting through. Getting through."

Teacher Expectations

Educator 1 explained how teacher expectations effect grades in their classroom. They maintained that they had to be very firm on the expectations. If they were, they said students would begin to understand and meet the expectations. "Well, once I blasted them so many times, guess what they started doing? Paying attention to the rules." They added, "I mean, if I walk

around, and I'm like, 'Do it again. No, ma'am. No, sir, I don't think that's okay.' No, I just blast people, and finally, they just fall in line and everybody starts doing it."

Educator 2 reflected on the collaboration among their department and how that affected grading. They shared that as a department they "talk a lot about grading and expectations, just so we make sure that we're being as consistent as possible. They added, "The purpose of grades is to let a student know where they are in relation to the expectations of the course, and, so, uh, that expectation should be grade level. Should be that."

Educator 3 articulated that they felt teachers had different expectations at different schools. Referencing students who transfer into Addie LaRue High School from other schools, they remarked, "You know, I...so, kids come to Addie LaRue High School, and they'll have a B. Well, I put them in class and start, and they're not level B. Where? Where did this come from?"

Reliability and Validity in Grading Practices

This theme builds upon the previously mentioned theme of teacher expectations. When grades are reliable and valid, they consistently and accurately measure what they intend to, and this is the same for all students. This would predict students with the same letter grade had performed very comparable to each other.

While this response from Educator 1 is linked to effort, it also indicated a difference in grading practices for English as a Second Language (ESL) and non-ESL students. There is also indication there are different grading practices among ability levels.

Educator 1 shared:

They give me kids who, they hate all teachers cause I can get them to pass and cause I'll get them to do work for me. And they give me ESL kids. Cause you know what? If they try, if they write some stuff...Cool. Great.

Educator 2 spoke to varying teacher expectations for students who are in different ability-leveled courses when asked about the purpose of grades. Prior to this quote, they were discussing the purpose of grades, and they had said grades should let students know where they are in relationship to the course requirements. They said:

...should be that...can kind of vary a little bit of, you know, the high school we are...you know, we, like, hyper level, so you know trying to figure out, you know, so a student that's passing my SpEd geometry class will have a different level of competency than a student that's passing an honors geometry class or even the next level up. But how close are they to expectations?

Educator 3 also noted variation in grading practices to determine grades. They revealed that even though the department had a practice for grading 60/20/20, this could be impacted by how teachers could tweak that policy by counting certain assignments more or less within the categories. They revealed, "And then you throw your own little percentages in there. Whatever you think it's going to be."

Summary

From this analysis, nine different themes emerged from the analysis of the data. The emergent themes associated with research question one were communication among stakeholders, student effort, and value of learning. From research question two, the theme of how to grade and implement grading practices emerged. This theme also emerged from research question three. The emergent theme from research question four was external pressures. Across

the research questions, the themes of grades as motivators, teacher expectations, and reliability and validity in grading practices emerged. Chapter 5 summarizes the findings, includes implications of the research, and recommendations for practice and future research.

Chapter 5. Conclusions

This chapter includes the findings, conclusions, and recommendations for this study. The purpose of the phenomenological study was to understand the educator experience of choosing grading practices for math and English courses East Tennessee high schools. Four research questions were developed to understand the essential question, and the researcher developed nine open-ended interview questions to support the four research questions.

To understand the educator experience, the researcher developed the essential research question to guide the study: What are the lived experiences of high school math and English teachers when choosing grading practices in their classrooms?

The following research questions supported the essential question of the research:

RQ1: How do teachers perceive the use of grades in secondary math and English classrooms?

RQ2: What experiences are associated with choosing a classroom grading practice for secondary math and English classrooms?

RQ3: What experiences have teachers had with professional development centered on grading practices and policy for the classroom?

RQ4: What external influences affect teacher choice in classroom grading practices?

Bandura's Social Learning Theory and Social Cognitive Theory were used as the framework to understand how teachers are influenced by their own beliefs about grading, the impact of their educational experience and other teachers, and the culture of grading in their current school.

Discussion

Research Question One

How do teachers perceive the use of grades in secondary math and English classrooms?

The themes of communication among stakeholders, student effort, and value of learning emerged from this research question.

Communication Among Stakeholders. All the educators indicated that grades are used as a form of communication among stakeholders. This included both as feedback on assignments to the students and as formal communication to stakeholders regarding the student's level of proficiency.

The literature indicates grades are an accepted norm in education and a professional responsibility of educators (Adom et al., 2020; Schneider & Huff, 2013). While teachers were able to personally visit students' homes to communicate student progress in the early days of education when fewer students attended school, stakeholders now depend on grades to communicate these measures (Brookhart et al., 2016; Guskey, 2015; Schneider & Huff, 2013).

Student Effort. A second theme for this question was student effort. For one educator, they felt effort, and in some cases attitude, should impact and possibly determine the student's grade. While one educator spoke to the fairness of the weighted grading system, they also added the homework category was more about grading behavior. The last educator believes grading homework for accuracy versus completion would increase the student's effort. These responses evoked the image of grades as currency to influence either academic and/or personal behaviors.

The literature indicates student effort has erroneously been tied to grades. Wormeli (2006) and Brookhart et al. (2016) both found educators believe effort should be included in grades. Cross and Frary (1996) along with Link and Guskey (2019) both found grades are a combination of both assessment and effort, which is not recommended best practice (Stiggins et al., 1989). Combining these prevents a clear understanding of the grade and can mislead

stakeholders about the student's proficiency (Guskey, 2006). Yeh (2019) posited this inclusion of non-academic factors unfairly contributes to the achievement gap.

Value of Learning. The final theme for this research question was the value of learning. While this emerged as a theme from all educators, it was from a slightly different perspective from each. One educator revealed their support for students to redo work because they want to focus on learning and not a grade, and another spoke to a student choosing not to complete all his homework because he felt he understood the material well enough. The third educator referenced their previous belief that completing homework for accuracy instead of completion would increase student learning. An additional layer to this theme is educators grading differently based on what they value personally. Although this theme presented itself through different lenses, each educator conveyed that a certain value is placed on learning by teachers and students.

The literature indicates the value placed on learning is affected by several factors. Dweck (2016) and Nottingham and Larsson (2022) proposed students' mindsets affect their attitudes toward learning and their ability to learn depending on if it is a fixed or growth mindset. Dweck (2016) and Nottingham and Larsson (2022) also posited an educator's fixed or growth mindset affects teachers' bias toward students. For example, an educator's fixed, negative mindset might be a bias toward a student that diminishes the support that student receives. Yeager et al. (2019) theorized the value students place on learning correlates to the consequences they face either academically or socially if they fail. If a student's mindset is that failure is too great a risk, they will place a low value on learning, but if they value learning, they will accept this risk (Yeager et al., 2019).

Theoretical Framework

The three themes that emerged from this research question were communication among stakeholders, student effort, and value of learning. When grades are communicated to students and/or other stakeholders, they can create a change in the student's behavior, especially if the grade was lower than expected. It is also a signal to the others to either continue or change the behavior. For example, a teacher who has successful students may continue to teach a certain course, but a teacher whose students are failing may be moved to a different course. Additionally, teachers may receive this as positive feedback if students' grades are good in their opinion to continue the behavior (teaching) they exhibited.

When students receive grades they are happy with, this is categorized as a positive response, and they will continue the behavior that helped them get this grade. Value of learning is also an exchange among the environment, cognition, and behavior. In this study, one teacher expressed their belief that correcting work helped students. They also indicated they were successful in getting student to learn. This belief is reinforced by the confidence administrators have placed in them to support students who have struggled with other teachers.

Research Question Two

What experiences are associated with choosing a classroom grading practice for secondary math and English classrooms?

From this question, the theme of how to grade and implement grading practices emerged. Through their responses, the researcher found all participants had minimal experience with grades and grading practices in their preservice education, and they often adapted their own experiences as a student with what they learned from colleagues.

One educator shared they had only learned how to develop a rubric, which is not appropriate for every grading situation. Another educator learned how to develop a test question, but it was not clear how that information was to be used in relationship to a grade for a marking period or what other factors should be included in a marking period. The third educator assumed grading was an understood practice.

The literature indicates teachers do not receive adequate training on grades and how to implement grading practices. Guskey (2009) and Lambert (1989) observed that many educators do not receive any formal training for grading, so they have an insufficient understanding about grades and/or reporting them to stakeholders. Bonner and Chen (2009) found that preservice grading positions did change once educators began teaching, and they suggested this might be because educators learn more about validity in relationship to grading.

Theoretical Framework. How to grade and implement grading practices emerged from this research question. In this question, educators shared their preservice experience with grades and grading practices. Each attested they had little to no experience, and grades were a mixture of what they had experienced as a student (personal belief) and what they had learned from mentors or colleagues (environment). Once they began practicing, they began to modify their initial practice (behavior) based on feedback. Eventually they develop a practice that remains in practice until some force compels them to change.

Research Question Three

What experiences have teachers had with professional development centered on grading practices and policy for the classroom?

While this question was similar to research question two, the objective was to determine if there was a difference in the level of importance placed on grades and grading practices once

educators entered the classroom. The theme of how to grade and implement grading practices emerged again from this question.

Each of the educators shared they had had minimal opportunities to learn more about grades and grading practices. Two of them indicated they had spent some time learning about standards-based grading, but the third answered they had no knowledge of any professional development around grades or grading practices. Standards-based grading was an initiative of previous administration at Addie LaRue High School, and it is possible specific educators were intentionally chosen to attend professional development for this. It is possible that the opportunity was not available to everyone.

Similar to research question two, Guskey (2009) found there is a lack of formal training for grading practices, so educators are often using an assortment of practices to create their own (Guskey, 2006; McMillan, 2001). This approach to grading creates concerns for validity in grading (Guskey, 2015) and confusion about the meaning of the grade for stakeholders (Brookhart et al., 2016; Tomlinson, 2005). Although Bonner and Chen (2009) found preservice grading positions changed once educators were practicing in the classroom, the new grading practice could still be inappropriate due to minimal professional development opportunities or professional learning communities.

Theoretical Framework. How to grade and implement grading practices also emerged from this research question. In this question, educators shared their professional development experiences as practicing educators. As indicated in question three, once a grading practice is in place, it will remain in place until there is some force or feedback to change it. The longer it goes unchanged, the more fixed it becomes, which can make the educator more resistant to change. This can make professional development to enact change more challenging.

Research Question Four

What external influences affect teacher choice in classroom grading practices?

The theme of external pressures emerged from this research question. The interview questions inquired about the culture of grades and grading practices at their school and asked the educators to share their experiences with external influences. While the first interview question asked the educators about the culture of grading in their school, the following interview question to support this research question inquired about external factors impacting grades and/or grading practices. All three educators indicated there was an external pressure in some form that placed demands on students and educators in relation to grades and grading practices.

Two of the three educators expressed they had been pressured by parents to somehow modify their practice to improve a student's grade. Educator 1 theorized that because grades were all parents cared about, it was all the students cared about, too. For Educator 1, who previously taught Honors English for 15 years at another school, this led to them moving to a lower ability level course because of the pressure and expectations of honor students' parents that their students to make only As in her class.

Educator 2 spoke to the pressure placed on students and parents and their fear of being considered a failure if they make less than an A. They recounted their experience with being accused by a parent for ruining the student's life when he received a bad grade in the course.

Educator 3 said the only time students feared a bad grade was when there might be parent backlash, but their experience was that most parents did not care. Educator 3 reflected that the pressure is also placed on the school and school system by the state to have high graduation rates, which trickles down to students, staff, and parents.

The literature indicates parents pressure teachers to alter grades or change expectations for their children because of the competition in admissions to schools and for scholarship monies (Lloyd, 2019). This pressure can lead to grade inflation (Tyner & Gershenson, 2020). There is also pressure from schools to not fail students due to the increased costs of a student repeating a course (Gonzalez, 1995). Feldman (2020) noted that a stakeholder's disagreement with a grade was fundamentally the questioning of the educator's professional expertise, which is a source of stress for educators.

Theoretical Framework. This theme, external pressure, emerged when educators were asked about the culture of grades and grading practices at their school and what, if any, external pressures they encountered. These pressures included pressure from parents and from the state level. Through the framework, these external pressures impact both students and teachers deeply. Educator 1 spoke to their desire to no longer teach honors courses because of the pressure from parents. In that example, their behavior (teaching) was impacted by the environment (parents). Educators 2 and 3 shared examples where either the pressure from a parent or from the state impacted grades and grading practices. In this case, beliefs (grades less than A are unacceptable; schools must have high graduation rates at all costs) impact the behaviors of educators. In some cases, it could cause schools to weaken the rigor of the coursework to meet these beliefs.

Additional Themes

While analyzing the data, three themes emerged that were not associated with a specific research question but appeared across more than one research question. These themes are grades as motivators, teacher expectations, and reliability and validity in grading practices.

Grades as Motivators. The educators revealed that grades are indeed motivators for students. Two educators indicated motivation to attain high grades, but one educator indicated

motivation to be placed in an easier online program was a motivator for lower or failing grades. One educator tied this back to external pressures by relating that some students are motivated because they feel pressured by their parents.

The literature indicates that prior research has also shown grades are used by educators to promote certain preferred behaviors or compliance with classroom rules (Chiekem, 2015; Vatterott, 2015). Guskey (2009) found these practices to be more common in secondary classrooms than elementary. It is most often seen in assigning zeros for missing work or deducting points for late work or behaviors the educator finds unacceptable (Guskey, 2009; Hochweber et al., 2014). Students who are motivated by grades will typically meet the behavior expectations as part of attaining their desired grade.

Teacher Expectations. Each educator spoke to the role of teachers and the theme of teacher expectations emerged. They all noted that the teacher expectation for work influenced the performance of the student. This was true in of both academics and behavior. Educator 3 noted that not all teachers, departments, or schools had the same academic expectations for students.

The literature indicates teacher expectations of students impacts student performance. Dueck (2020) advanced the theory that a student's belief about themselves is an important factor in their academic success, and this is influenced by their past experiences with grades. McFarland et al. (2016) found students are significantly affected by how they feel their teachers perceive them as a student. Teacher perception of student performance is a bias in grading (Andersen, 2018), however, and teachers may grade students unfairly based on their expectations of the student (Hardré, 2018). Zoeckler (2007) also found that educator opinions of students, such as careless or diligent, were incorporated into grades despite the participants reporting that student character should not be included in grades.

Reliability and Validity in Grading Practices. Responses across the interviews indicated the educators adjusted their grading practices and/or expectations when they felt it was prudent. Educator 1 shared their experience with teaching ESL students and students who held a fixed, negative mindset toward other teachers, while Educator 2 disclosed various levels of expectations for different ability levels of students. Educator 3 referenced influencing percentages with the department's expectations for weighted grades, such as counting a test grade twice.

The literature indicates a history of issues with reliability and validity related to grading practices. Brookhart et al. (2016) noted grades and grading practices have been questioned for over 100 years and found there was noteworthy diversity in teachers' grading practices that created issues with reliability and validity in grades. Gershenson (2020) observed that assigning grades that do not show signs of inflation or inaccuracy had a positive impact on Black and Hispanic students' learning, while Link and Guskey (2019) postulated that deducting points for cultural differences in expectations was a form of inaccuracy in grading. Nottingham and Larsson (2022) described in their study how an educator's mindset about a student can alter the amount of support that student receives. The literature notes a myriad of ways the reliability and validity of grades are affected by the assortment of grading practices.

Theoretical Framework. The themes of grades as motivators, teacher expectations, and reliability and validity in grading practices were woven into all the research questions. When exploring grades as motivators, the behavior is impacted by response received. If a student believes studying for a test will be rewarded with a good grade and that is what happens, they will continue to believe that is true until proven otherwise, at which time they may not change their mind but might re-evaluate the belief. When teachers have different expectations for

different students, and the environment reinforces that idea by accepting and/or encouraging it, that behavior will continue. Regarding reliability and validity in grading, when an educator holds a particular belief about this and the environment continues to support the behavior, it will not change until the belief or the behavior is challenged.

Implications for Practice

Based on the analysis of the data and the literature review in Chapter 2, the following recommendations were made after considering the experiences of educators when selecting a grading practice. These recommendations are intended to support future educators as they choose a grading practice for their classrooms by providing a foundation of understanding for the reliability and validity of grades and grading practices and to encourage a positive grading culture for all stakeholders.

The first implication for practice is to meet the need for additional preservice coursework regarding reliability and validity in grades and grading practices. Although there was no analysis of any university coursework as part of this study, the research indicated a need to incorporate measurement courses along with the pedagogical coursework.

The second implication for practice is to encourage ongoing professional development and/or professional learning communities for practicing educators. It is recommended that educators become advocates for the change they wish to see in grades, grading practices, and culture of grading. This can be accomplished through formal outside professional development, book studies within the school or a department, or focused, intentional conversations to consider school policy and practice in grades and grading practices. Teachers should reflect on if the grading practices are accurate measure of student achievement.

The third implication for practice is to increase support at the school and district levels for teachers and students in relation to external pressures. Supports could include incorporating mindset/SEL curriculum for all students; mental health counseling for those in need of additional support; and increasing conversations with stakeholders to shift the focus from grades to learning.

School administration and counselors may also focus on informing students and stakeholders about the complete body of options available after high school graduation, such as military service, trade schools, community college, universities, or entering the workforce. Normalizing a variety of options may help to alleviate the pressures currently surrounding grades and grading practices.

Recommendations for Future Research

The literature reviewed in Chapter 2 and the results of this study both indicate there is a lack of accuracy in grades and grading practices in high school math and English courses. This can be attributed to several factors: lack of preservice coursework in measurement, inadequate professional development for practicing teachers, and external impacts on grades and grading practices. Based on this, three recommendations are made for future research.

Firstly, it is recommended that the study be expanded to include educators from other content areas. Math and English grading practices are more opposite than science or social studies might be, so different or additional themes may emerge when these content areas are incorporated. This additional data could narrow future recommendations to be more effective in creating change in grading practices.

The second recommendation is to replicate this study and include students' and parents' experiences with grades to learn more about the starting point of the external pressure discussed

in this study. Identifying the source and tracing its evolution may help develop strategies and communication to prevent external pressure from developing.

The final recommendation is for future research to ascertain what current required courses are in place at colleges and universities so that students learn the importance of accuracy in grades and grading practices and understand how to make a grade as reliable and valid as possible before becoming practicing educators. This recommendation is made to increase accuracy in grades and grading practices.

Summary

In this chapter, the findings from Chapter 4 were discussed along with the implications for practice. Bandura's Social Learning Theory and Social Cognitive Theory were used as the theoretical framework for this. Recommendations for future study were also made based on the findings. These recommendations include expanding the study to include other content areas, replicating the study to include parents and students, and investigating preservice coursework related to measurement in grades and grading practices.

REFERENCES

- ACT. (2023). *The ACT test overview*. <https://www.act.org/content/act/en/products-and-services/the-act.html>
- Adom, D., Adu Mensah, J., & Dake, D.A. (2020). Test, measurement, and evaluation: Understanding and use of the concepts in education. *International Journal of Evaluation and Research in Education*, 9(1), p. 109-119. <https://doi.org/10.11591/ijere.v9i1.20457>
- Adu-Mensah, J. (2018). The attitude of basic schoolteachers toward grading practices: Developing a standardized instrument. *International Journal of Social Sciences & Educational Studies*, 5(1), 52-62. <https://doi.org/10.23918/ijsses.v5i1p52>
- Ahmed, F., Ali, S., & Shah, R.A. (2019). Exploring variation in summative assessment: Language teachers' knowledge of students' formative assessment and its effect on their summative assessment. *Bulletin of Education and Research*, 41(2), p. 109-119. <https://files.eric.ed.gov/fulltext/EJ1229441.pdf>
- Allen, J.D., & Lambating, J. (2001, April 10-14). *Validity and reliability in assessment and grading: Perspectives of preservice and inservice teachers and teacher education professors* [Paper presentation]. American Educational Research Association Annual Meeting, Seattle, WA, United States.
- Andersen, I.G. (2018). Pygmalion in instruction? Tracking, teacher reward structures, and educational inequality. *Social Psychology of Education*, 2018(21), 1021-1044. <https://doi.org/10.1007/s11218-018-9452-z>
- Anderson, L.W. (2018). A critique of grading: Policies, practices, and technical matters. *Education Policy Analysis Archives*, 26(49), p. 1-27.

- Ansary, T. (2007, March 9). *Education at risk: Fallout from a flawed report*. Edutopia.
<https://www.edutopia.org/landmark-education-report-nation-risk>
- Ansell, S.E. (2011, July 7). *Achievement gap*. Education Week. Retrieved July 12, 2022 from
<https://www.edweek.org/leadership/achievement-gap/2004/09>
- Amankwaa L. (2016). Creating protocols for trustworthiness in qualitative research. *Journal of Cultural Diversity*, 23(3), 121–127.
- Arnsten, A.F.T. (2009). Stress signalling pathways that impair prefrontal cortex structure and function. *Nature Reviews Neuroscience*, 10, p. 410–422.
<https://doi.org/10.1038/nrn2648>
- Azzam, A.M. (2007). Why students drop out. *Educational Leadership*, 64(7), 91-93.
- Bandura, A. (1977). *Social Learning Theory*. Prentice-Hall, Inc.
- Bandura, A. (1986). *Social Foundations of Thoughts & Actions: A Social Cognitive Theory*. Prentice-Hall, Inc.
- Bonner, Sarah & Chen, Peggy. (2009). Teacher candidates' perceptions about grading and constructivist teaching. *Educational Assessment*, 14, p. 57-77.
<https://doi.org/10.1080/10627190903039411>
- Bouchrika, I. (2021, August 9). *Standards-based grading: Definition, benefits, & comparison with traditional grading*. <https://research.com/education/standards-based-grading>
- Brimi, H. (2011). Reliability of grading high school work in English. *Practical Assessment, Research, and Evaluation*, 16(17). <https://doi.org/10.7275/j531-fz38>

- Brookhart, S.M. (1993). Teachers' grading practices: Meaning and values. *Journal of Educational Measurement*, 30(2), p. 123-142.
- Brookhart, S.M., Guskey, T.R., Bowers, A.J., McMillan, J.H., Smith, J.K., Smith, L.F., Stevens, M.T., & Welsh, M.E. (2016). A century of grading research: Meaning and value in the most common educational measure. *Review of Educational Research*, 86(4), p. 803-848.
- Burke, K. (2010). *Balanced Assessment*. Solution Tree Press.
- Burleigh, T.J., & Meegan, D.V. (2018). Risky prospects and risk aversion tendencies: Does competition in the classroom depend on grading practices and knowledge of peer-status? *Social Psychology of Education*, 21(2018), p. 323-335. <https://doi.org/10.1007/s11218-017-9414-x>
- Carnegie Mellon University Eberly Center. (2023). *What is the difference between formative and summative assessment?* <https://www.cmu.edu/teaching/assessment/basics/formative-summative.html>
- Chen, P.P., and Bonner, S.M. (2017). Teachers' beliefs about grading practices and a constructivist approach to teaching. *Education Assessment*, 22(1), p. 18-34. <https://dx.doi.org/10.1080/10627197.2016.1271703>
- Chiekem, E. (2015). Grading practice as valid measures of academic achievement of secondary schools students for national development. *Journal of Education and Practice*, 6(26), p. 24-29. <https://files.eric.ed.gov/fulltext/EJ1077389.pdf>
- Conant, J. B. (1959). *The American high school today: A first report to interested citizens; The American high school today: A first report to interested citizens*. McGraw-Hill Book Company. <https://doi.org/10.1037/13171-003>

- Cooper, H., Robinson, J. C., & Patall, E. A. (2006). Does homework improve academic achievement? A synthesis of research, 1987–2003. *Review of Educational Research*, 76(1), 1–62. <https://doi.org/10.3102/00346543076001001>
- Copa, G.H., & Pease, V.H. (1992). *The comprehensive high school: An historical perspective*. National Center for Research in Vocational Education. <https://files.eric.ed.gov/fulltext/ED352520.pdf>
- Creswell, J.W., & Creswell, J.D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). Sage Publications.
- Creswell, J.W., & Poth, C.N. (2018). *Qualitative inquiry & research design: Choosing among five approaches* (4th ed.). Sage Publications.
- Creswell, J.W., Hanson, W.E., Plano Clark, V.L., & Morales, A. (2007). Qualitative research designs: Selection and implementation. *The Counseling Psychologist*, 35(2), p. 236-264. <https://doi.org/10.1177/0011000006287390>
- Crofton, R. (2020, May 15). *That old deluder satan: Puritan emphasis on compulsory education*. The Paul Revere House. <https://www.paulreverehouse.org/that-old-deluder-satan-puritan-emphasis-on-compulsory-education/>
- Cross, L.H., & Frary, R.B. (1996, April 9-11). *Hodgepodge grading: Endorsed by students and teachers alike* [Paper presentation]. National Council on Measurement in Education Annual Meeting, New York, NY, United States.
- Dalton, B., Glennie, E., & Ingels, S. (2009). *Late high school dropouts: Characteristics, experiences, and changes across cohorts*. National Center for Education Statistics, U.S. Department of Education. <https://nces.ed.gov/pubs2009/2009307.pdf>

- Darlaston-jones, D. (2007). Making connections: The relationship between epistemology and research methods. *Australian Community Psychologist*, 19(1), 19-27.
- Donovan, C. (2017, August 19). *Classical Test Theory* [Video]. YouTube.
<https://youtu.be/TLIPGFT952M>
- Doornkamp, L., Van der Pol, L.D., Groeneveld, S., Mesman, J., Endendijk, J.J., Groeneveld, M.G. (2022, August 6). Understanding gender bias in teachers' grading: The role of stereotypical beliefs. *Teaching and Teacher Education*, 118(2022), 1-14.
<https://doi.org/10.1016/j.tate.2022.103826>
- Dueck, M. (2020, September 18). *Stepping up the grading environment*. Educational Leadership.
<https://www.ascd.org/el/articles/stepping-up-the-grading-environment>
- Duke, D.L., & Canaday, R.L. (1991). *School Policy*. McGraw-Hill, Inc.
- Duplechain, R., Reigner, R., & Packard, A. (2008). Striking differences: The impact of moderate and high trauma on reading achievement. *Reading Psychology*, 29(2), p. 117-136.
https://www.tandfonline.com/doi/pdf/10.1080/02702710801963845?casa_token=3bX1f-Rl2f8AAAAA:lnDFrqWomLZCJbgPH_SAQ2khej6sUxsD3GTne4wK49C3pIGPmZPB UPTYCAcszmOvHQm9BKI_TCA
- Durm, M.W. (1993). An A is not an A is not an A: A history of grading. *The Educational Forum*, 57(Spring).
https://www.spsd.org/uploaded/PDFs/Curriculum/PbL/Durm_1993_An_A_Is_Noit_an_A_.pdf
- Dweck, C.S. (2016). *Mindset: The new psychology of success*. Ballentine Books.

- Edglossary. (2014, February 4). *Learning Standards*. <https://www.edglossary.org/learning-standards/>
- Edvocate. (n.d.). *A nation at risk*. <https://www.theedadvocate.org/edupedia/content/what-was-a-nation-at-risk-the-imperative-for-educational-reform/>
- Edwards, H. (2021, February 25). *What are AP classes? Why should you take them?*
PrepScholar. [What Are AP Classes? Why Should You Take Them? \(prepscholar.com\)](https://prepscholar.com/what-are-ap-classes-why-should-you-take-them/)
- Epstein, J.L., & Van Voorhis, F.L. (2001). More than minutes: Teachers' roles in designing homework. *Educational Psychologist*, 36(3), p. 181-193.
- Eun, B. (2019). Adopting a stance: Bandura and Vygotsky on professional development. *Research in Education*, 105(1), p. 74-88. <https://doi.org/10.1177/0034523718793431>
- Feldman, J. (2020). Taking the stress out of grading. *Educational Leadership*, 78(1), 14-20.
<http://68.77.48.18/RandD/Educational%20Leadership/Grading-2020-09/Taking%20the%20Stress%20Out%20of%20Grading%20-%20Educational%20Leadership.pdf>
- Feldman, J. (2019). Beyond standards-based grading: Why equity must be part of grading reform. *Phi Delta Kappan*, 100(8), 52-55.
https://journals.sagepub.com/doi/pdf/10.1177/0031721719846890?casa_token=IU8MtpbW_mwAAAAA:R5SkT9v-1i0YgVzS6BJYF3ZL4HNMTpniWKnenACDb-74rJQkS_1kckzM31hqr9_N8QOE4NeVZak

Ferrari, M., Robinson, D.K., Tasnitsky, A. (2010). Wundt, Vygotsky and Bandura: A cultural-historical science of consciousness in three acts. *History of the Human Sciences*, 23(3), p. 95-119. <https://doi.org/10.1177/0952695110363643>

Films Media Group. (2008). *Formative assessment in content areas: High school. Films On Demand*. Retrieved February 16, 2023, from <https://fod.infobase.com/PortalPlaylists.aspx?wID=104058&xtid=204178>

Fitzner, K. (2007). Reliability and validity: A quick review. *The Diabetes Educator*, 33(5), p. 775-780.

Gershenson, S. (2020, February 4). *Great expectations: The impact of rigorous grading practices on student achievement*. Thomas B. Fordham Institute. <https://fordhaminstitute.org/national/research/great-expectations-impact-rigorous-grading-practices-student-achievement>

Gershenson, S., & Tekin, E. (2018). The effect of community traumatic events on student achievement: Evidence from the Beltway Sniper attacks. *NBER Working Paper Series*, 21055-. <https://doi.org/10.3386/w21055>

Geving, A.M. (2007). Identifying the types of student and teacher behaviours associated with teacher stress. *Teaching and Teacher Education*, 23(5), p. 624-640. <https://doi.org/10.1016/j.tate.2007.02.006>

Giota, J. & Gustafsson, J. (2021). Perceived academic demands, peer and teacher relationships, stress, anxiety and mental health: Changes from grade 6 to 9 as a function of gender and cognitive ability. *Scandinavian Journal of Educational Research*, 65(6), p. 956-971. <https://doi.org/10.1080/00313831.2020.1788144>

- Gonzalez, D. (1995, July 3). In schools, subtle pressure to change students' grades. *The New York Times*. <https://www.nytimes.com/1995/07/03/nyregion/in-schools-subtle-pressure-to-change-students-grades.html>
- Green, S.K., Johnson, R.L., Kim, D., Pope, N.S. (2007). Ethics in classroom assessment practices: Issues and attitudes. *Teaching and Teacher Education*, 23(2007), 999-1011. <https://doi.org/10.1016/j.tate.2006.04.042>
- Greeneville City Schools. (n.d.). *Core Values, Mission & Vision*. https://www2.gcschools.net/apps/pages/index.jsp?uREC_ID=1131446&type=d&pREC_ID=1401140
- Greeneville City Schools Education Foundation [@gcs.edu.foundation]. (2020, April 22). *Congratulations to the GHS ACT 30+ Club for 2019-20! You can find these smiling faces on the Tusculum Blvd digital billboard. @PatrickFraleyg @GreenevilleCity @writenoelle @GHS_MainOffice @steve_starnes* [image attached] [Tweet]. Twitter. <https://twitter.com/gcsedufoundatn/status/1253040660166389760>
- Greeneville High School. (2022). *Student Policy Handbook 2022-23*. <https://4.files.edl.io/10a5/09/09/22/155827-f576d49e-2c5c-4e0a-99e9-e56f4ea6e69b.pdf>
- Griffin, R. & Townsley, M. (2021). Point, points, and more points: High school grade inflation and deflation when homework and employability scores are included. *Journal of School Administration Research and Development*, 6(1), p. 1-11. <https://files.eric.ed.gov/fulltext/EJ1301287.pdf>
- Guskey, T.R. (2006). Making high school grades meaningful. *Phi Delta Kappan*, 87(9), p. 670-675. <https://doi.org/10.1177%2F003172170608700910>

- Guskey, T.R. (2009, April). *Bound by tradition: Teachers' views of crucial grading and reporting issues* [Paper presentation]. Annual Meeting of the American Educational Research Association, San Francisco, CA, United States.
- Guskey, T.R. (2013). The case against percentage grades. *Educational, School, and Counseling Psychology Faculty Publications*, 71(1), p. 68-72.
https://uknowledge.uky.edu/edp_facpub/22
- Guskey, T.R. (2015). *On your mark: Challenging the conventions of grading and reporting*. Solution Tree Press.
- Guskey, T. R., & Bailey, J. M. (2010). *Developing Standards-Based Report Cards*. United States: SAGE Publications.
- Guskey, T.R., & Brookhart, S.M. (Eds.). (2019). *What we know about grading: What works, what doesn't, and what's next*. ASCD.
- Guskey, T.R., Townsley, M., & Buckmiller, T.M. (2020). The impact of standards-based learning: Tracking high school students' transition to the university. *NASSP Bulletin*, 104(4), p. 257-269. <https://doi.org/10.1177/0192636520975862>
- Guttek, G.L. (1991). *Education in the United States: An historical perspective*. Allyn & Bacon.
- Guy-Evans, O. (2021, May 8). Frontal lobe: Function, location, and structure. SimplyPsychology. <https://www.simplypsychology.org/frontal-lobe.html>
- Haladyna, T.M. (2019, May). *Assigning a valid and reliable grade in a course*. IDEA. <https://files.eric.ed.gov/fulltext/ED598950.pdf>

Hall, M. (2013, May 13). *To curve or not to curve*. The Innovative Instructor.

<https://ii.library.jhu.edu/tag/criterion-referenced-grading/>

Hardré, P. (2014). Checked your bias lately. Reasons and strategies for rural teachers to self-

assess for grading bias. *Rural Educator*, 35(2). <https://doi.org/10.35608/ruraled.v35i2.352>

Harlen, W., & James, M. (1997). Assessment and learning: differences and relationships between

formative and summative assessment. *Assessment in Education: Principles, Policy &*

Practice, 4(3), p. 365-379. <https://dx.doi.org/10.1080/0969594970040304>

Harvard Health Publishing. (n.d.). Understanding the stress response.

<https://www.health.harvard.edu/staying-healthy/understanding-the-stress-response>

Hochweber, J., Hosenfeld, I., & Klieme, E. (2014). Classroom composition, classroom

management, and the relationship between student attributes and grades. *Journal of*

Educational Psychology, 106(1), 289-300. <https://doi.org/10.1037/a0033829>

Högberg, B., Lindgren, J., Johansson, K., Strandh, M., & Petersen, S. (2021) Consequences of

school grading systems on adolescent health: Evidence from a Swedish school reform.

Journal of Education Policy, 36(1), 84-106.

<https://doi.org/10.1080/02680939.2019.1686540>

Hurst, D., Tan, A., Meek, A., & Sellers, J. (2003). *Overview and inventory of state education*

reforms: 1990 to 2000. U.S. Department of Education National Center for Education

Statistics (2003). <https://nces.ed.gov/pubs2003/2003020.pdf>

Iachini, A.L., Petiwala, A.F., & DeHart, D.D. (2016). Examining adverse childhood experiences

among students repeating ninth grade: Implications for school dropout prevention.

Children & Schools, 38(4), p. 218-227.

Iamarino, D. (2014, May 21). The benefits of standards-based grading: A critical evaluation of modern grading practices. *Current Issues in Education*, 17(2), p. 8-17.

Individuals with Disabilities Education Act, 20 U.S.C. § 300.320 (2004).

International Affairs Office (2008). *Structure of the U.S. education system: U.S. grading systems*. U.S. Department of Education.

<https://www2.ed.gov/about/offices/list/ous/international/usnei/us/grading.doc>

Issayeva, L. (2021, October 31). *What is test equating?* ASC. <https://assess.com/test-equating/>

Jimenez, M.E., Wade, R., Schwartz-Soicher, O., Lin, Y., & Reichman, N.E. (2017). Adverse childhood experiences and ADHD diagnosis at age 9 years in a national urban sample. *Academic Pediatrics*, 17(4), p. 356-361. <https://doi.org/10.1016/j.acap.2016.12.009>

Johnson, S. (2018, February 8). Survey shows pressure teachers face to change grades and attendance. *New York Amsterdam News*.

<https://amsterdamnews.com/news/2018/02/08/survey-shows-pressure-teachers-face-change-grades/>

Jones, C. (2022, March 29). Parents' guide to 504 plans and IEPs: What they are and how they're different. *EdSource*. <https://edsources.org/2022/parents-guide-to-504-plans-and-ieps-what-they-are-and-how-theyre-different/669493#:~:text=A%20504%20plan%20is%20geared,have%20one%20or%20the%20other.>

- Kaftan, J. M., Buck, G. A., & Haack, A. (2006). Using formative assessments to individualize instruction and promote learning. *Middle school Journal*, 37(4), 44-49.
<https://files.eric.ed.gov/fulltext/EJ752868.pdf>
- Kamanetz, A. (2018, April 29). *What 'a nation at risk' got wrong, and right, about U.S. schools*. NPR. <https://www.npr.org/sections/ed/2018/04/29/604986823/what-a-nation-at-risk-got-wrong-and-right-about-u-s-schools>
- Keane, G., & Heinz, M. (2019). Differentiated homework: Impact on student engagement. *Journal of Practitioner Research*, 4(2).
<https://digitalcommons.usf.edu/cgi/viewcontent.cgi?article=1111&context=jpr>
- Khan, S. (n.d.). *Statistics Intro: Mean, Median Mode*. Khan Academy.
<https://www.khanacademy.org/math/cc-sixth-grade-math/cc-6th-data-statistics/mean-and-median/v/statistics-intro-mean-median-and-mode>
- Kober, N., & Rentner, D.S. (2020). *History and evolution of public education in the U.S*. Center on Education Policy. <https://files.eric.ed.gov/fulltext/ED606970.pdf>
- Kohn, A. (2006). *The Homework Myth*. Lifelong Books.
- Kridel, C. (2010). *Cardinal principles of secondary education*. In *Encyclopedia of curriculum studies* (Vol. 1, pp. 98-98). SAGE Publications, Inc.,
<https://dx.doi.org/10.4135/9781412958806.n54>
- Lambert, N.M. (1989, August 11-15). *The crisis in measurement literacy in psychology and education* [Paper presentation]. Annual Meeting of the American Psychological Association, New Orleans, LA, United States.
- Leedy, P.D., & Ormrod, J.E. (2019). *Practical research* (12th ed.). Pearson.

- Library of Congress (n.d.). *America at School*. <https://www.loc.gov/collections/america-at-work-and-leisure-1894-to-1915/articles-and-essays/america-at-school/>
- Link, L.J., & Guskey, T.R. (2019). How traditional grading contribute to student inequities and how to fix it. *Educational, School, and Counseling Psychology Faculty Publications*, 53. https://uknowledge.uky.edu/edp_facpub/53
- Lloyd, C. (2019, May 13). *Grade grubbing: when parents cross the line*. Great Schools. <https://www.greatschools.org/gk/articles/grade-grubbing-parents/>
- Lynch, M. (2021, February 10). *A guide to criterion-referenced tests*. The Edvocate. <https://www.theedadvocate.org/criterion-referenced-test/>
- Lynch, M. (2016, November 18). *The pro's and cons of teacher-designed tests*. The Edvocate. <https://www.theedadvocate.org/pros-cons-teacher-designed-tests/#:~:text=Classroom%20assessments%20are%20a%20big,teachers%20design%20their%20own%20assessments.>
- Lyons, S.D., & Berge, Z.L. (2012). Social Learning Theory. In N.M. Seel, (Ed.) *Encyclopedia of the Sciences of Learning*. Springer, Boston, MA. https://doi.org/10.1007/978-1-4419-1428-6_1257
- McBride, A.E., & Kahle, A.J. (2018). Standards-based grading: an examination of assessment options. *Journal of Education & Social Policy*, 5(4), p. 140-145.
- McCombs School of Business (2018, December 18). *Morals/ ethics defined [Video]*. YouTube. <https://youtu.be/0WxOGR6HKFs>

- McFarland, L., Murray, E., & Phillipson, S. (2016). Student–teacher relationships and student self-concept: Relations with teacher and student gender. *Australian Journal of Education*, 60(1), p. 5-25. <https://doi.org/10.1177/0004944115626426>
- McMillan, J.H. (2001). Secondary teachers' classroom assessment and grading practices. *Educational Measurement: Issues and Practice*, 20(1), p. 20-32
- McMillan, J.H., & Nash, S. (2000, April 25-27). *Teacher classroom assessment and grading practices decision making* [Paper presentation]. National Council on Measurement in Education Annual Meeting, New Orleans, LA, United States.
- McMillan, J.H., & Workman, D.J. (1998). *Classroom assessment and grading practices: A review of the literature*. Metropolitan Educational Research Consortium, Virginia Commonwealth University, Richmond, VA.
- Malouff, J. (2008). Bias in grading. *College Teaching*, 56(3), 191–192.
<http://www.jstor.org/stable/20695206>
- Mangrum, M. (2022, April 17). *State testing counts again for students, districts. Here's what you should know*. The Tennessean.
<https://www.tennessean.com/story/news/education/2022/04/18/tennessee-state-testing-tcap-count-testing-begin/7273370001/>
- Marzano, R.J. (2017). *The new art and science of teaching*. ASCD/Solution Tree Press.
- Merriam, S.B. (1995). What can you tell from an N of 1?: Issues of validity and reliability in qualitative research. *PAACE Journal of Lifelong Learning*, 4, p. 51-60.

- Merriam-Webster. (n.d.). Bias. In *Merriam-Webster.com dictionary*. Retrieved July 17, 2022, from <https://www.merriam-webster.com/dictionary/bias>
- Miller, J.A. (1992, September 23). *Energy dept. lab's education report still on hold*. Education Week. <https://www.edweek.org/education/energy-dept-labs-education-report-still-on-hold/1992/09>
- Miller, M.D. (2019, October 30). *Reliability in educational assessments*. Oxford Bibliographies. <https://www.oxfordbibliographies.com/view/document/obo-9780199756810/obo-9780199756810-0228.xml?rsk=WNqmqKe&result=202>
- Mirel, J. (2006). The traditional high school. *Education Next*, 6(1). <https://www.educationnext.org/the-traditional-high-school/>
- National Center for Education Statistics. (1993). *120 years of American education: A statistical portrait*. <https://nces.ed.gov/pubs93/93442.pdf>
- National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform: A report to the Nation and the Secretary of Education*. United States Department of Education. <https://files.eric.ed.gov/fulltext/ED226006.pdf>
- Neer, M. (2018, July 3). *EDI and Hattie's visible learning*. Data Works Educational Research. <https://dataworks-ed.com/blog/2018/07/edi-hatties-visible-learning/>
- Nottingham, J., & Larsson, B. (2022). *Mindset in education. Challenging mindset: Why a growth mindset makes a difference in learning – and what to do when it doesn't*. SAGE Books. <https://dx.doi.org/10.4135/9781071872727>

- Olsen, B., & Buchanan, R. (2019). An investigation of teachers encouraged to reform grading practices in secondary schools. *American Educational Research Journal*, 56(5), p. 2004-2039. <https://doi.org/10.3102/0002831219841349>
- Öztürk Gübes, N. (2021). An investigation into weighting problem in norm-referenced grading system. *Eurasian Journal of Educational Research*, 93, p. 337–356.
- Park, J. (2004, September 10). *A nation at risk*. Education Week. <https://www.edweek.org/policy-politics/a-nation-at-risk/2004/09>
- Paterson, W. (2021, December 8). *From 1871 to 2021: A Short History of Education in the United States*. Buffalo State. <https://suny.buffalostate.edu/news/1871-2021-short-history-education-united-states>
- Patton, M.Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). SAGE Publishing, Inc.
- Porche, M.V., Fortuna, L.R., Lin, J., & Alegria, M. (2011). Childhood trauma and psychiatric disorders as correlates of school dropout in a national sample of young adults. *Child Development*, 82(3), p. 982-998.
- Pui, P., Yuen, B., & Goh, H. (2021) Using a criterion-referenced rubric to enhance student learning: a case study in a critical thinking and writing module. *Higher Education Research & Development*, 40(5), 1056-1069. <https://doi.org/10.1080/07294360.2020.1795811>
- Qu, S.Q., & Dumay, J. (2011). The qualitative research interview. *Qualitative Research in Accounting & Management*, 8(3), 238-264. <http://doi.org/10.1108/11766091111162070>

- Quinn, D. M. (2021). How to reduce racial bias in grading. *Education Next*, 21(1).
<https://www.proquest.com/scholarly-journals/how-reduce-racial-bias-grading/docview/2475144944/se-2>
- Radio Greeneville. (2019). *Over the past ten years, Greeneville has been given the nickname, "Titletown" for the athletic success of numerous teams throughout* [Image attached] [Status update]. Facebook.
<https://www.facebook.com/RadioGreenevilleNews/posts/over-the-past-ten-years-greeneville-has-been-given-the-nickname-titletown-for-th/1451517534996460/>
- Randall, J., & Engelhard, G. (2010). Examining the grading practices of teachers. *Teaching and Teacher Education*, 26(7), 1372-1380. <https://doi.org/10.1016/j.tate.2010.03.008>
- Reardon, S.F., Kalogrides, D., Fahle, E.M., Podolsky, A., & Zárate, R.C. (2018). The relationship between test item format and gender achievement gaps on math and ELA tests in fourth and eighth grades. *Editorial Researcher*, 47(5), 284-294. <http://edr.aera.net>
- Renaissance (n.d.). *What's the difference? Criterion-referenced tests vs. norm-referenced tests*.
<https://www.renaissance.com/2018/07/11/blog-criterion-referenced-tests-norm-referenced-tests/>
- Rosário, P., Núñez, J.C., Vallejo, G., Nunes, T., Cunha, J., Fuentes, S., & Valle, A. (2018). Homework purposes, homework behaviors, and academic achievement. Examining the mediating role of students' perceived homework quality. *Contemporary Educational Psychology*, 53, p. 168-180. <https://doi.org/10.1016/j.cedpsych.2018.04.001>
- Rossen, E. (2018, November). *Creating trauma-informed individualized education programs*. American Psychological Association.
<https://www.apa.org/pi/families/resources/newsletter/2018/11/trauma->

[teaching#:~:text=Rather%20than%20simply%20reducing%20a,identifying%20alternativ
e%20behaviors%20to%20interrupting.](#)

Rubel, L., & McCloskey, A. (2019). The "soft bigotry of low expectations" and its role in maintaining White supremacy through mathematics education. *Occasional Paper Series, 41*(10), 113-128. <https://educate.bankstreet.edu/occasional-paper-series/vol2019/iss41/10/>

Russo, C. J. (2008). *Stuart v. school district no. 1 of village of Kalamazoo*. *Encyclopedia of Education Law, 1*, p. 801-801. SAGE Publications, Inc.
<https://dx.doi.org/10.4135/9781412963916.n366>

Ryan, J. (2018). *A practitioner's introduction to equating with primers on classical test theory and item response theory*. Council of Chief State School Officers.
<https://ccsso.org/sites/default/files/2018-06/A%20Practitioners%20Introduction%20to%20Equating%20revised%20edition.pdf>

Ryan, R.M., & Deci, E.L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology, 61*(2020). <https://doi.org/10.1016/j.cedpsych.2020.101860>

Schneider, J., & Hutt, E. (2013, May 16). Making the grade: A history of the A-F marking scheme. *Journal of Curriculum Studies, 2013*.
<http://dx.doi.org/10.1080/00220272.2013.790480>

Schimmer, T. (2016). *Grading from the inside out: Bringing accuracy to student assessment through a standards-based mindset*. Solution Tree Press.

School Counseling Department. GHS Honors Program Requirements. Greeneville High School.

Retrieved February 26, 2023, from

https://ghs.gcschools.net/apps/pages/index.jsp?uREC_ID=1151664&type=d&pREC_ID=1511337

Schuster, C., Narciss, S., Bilz, J. (2021). Well done (for someone of your gender)! Experimental evidence of teachers' stereotype-based shifting standards for test grading and elaborated feedback. *Social Psychology of Education (2021)*24, p. 809-834.

<https://doi.org/10.1007/s11218-021-09633-y>

Scott, C. and Medaugh, M. (2017). Axial Coding. In J. Matthes, C.S. Davis, and R.F. Potter (Eds.), *The International Encyclopedia of Communication Research Methods*.

<https://doi.org/10.1002/9781118901731.iecrm0012>

Spears, A. (2015, June 15). *Quickly Inflated*. Tennessee Education Report.

<http://tnedreport.com/tag/cubed-root-method-and-quick-scores/>

Stanley, G., & Baines, L. (2001). No more shopping for grades at b-mart: Re-establishing grades as indicators of academic performance. *The Clearing House*, 74(4), p. 227–230.

<http://www.jstor.org/stable/30189667>

Starch, D., & Elliott, E.C. (1912). Reliability of the grading of high-school work in English. *The School Review*, 20(7), p. 442-457.

Starch, D., & Elliott, E.C. (1913). Reliability of grading work in mathematics. *The School Review*, 21(4), p. 254-259.

- Stauffer, B. (2022, May 10). 11 ways to prevent cheating in schools. *Applied Educational Systems*. <https://www.aeseducation.com/blog/ways-to-prevent-cheating-in-schools>
- Stiggins, R.J. (1993). Teacher training in assessment: overcoming the neglect. *Teacher Training in Measurement and Assessment Skills*, 4. Accessed 2/11/2023 from https://digitalcommons.unl.edu/burosteachertraining/4/?utm_source=digitalcommons.unl.edu%2Fburosteachertraining%2F4&utm_medium=PDF&utm_campaign=PDFCoverPages
- Stiggins, R.J. (2002). Assessment crisis: The absence of assessment for learning. *Phi Delta Kappan*, 83(10), p. 758–765. <https://doi.org/10.1177/003172170208301010>
- Stiggins, R.J., Frisbie, D.A. and Griswold, P.A. (1989). Inside high school grading practices: Building a research agenda. *Educational Measurement: Issues and Practice*, 8(2), p. 5-14. <https://doi.org/10.1111/j.1745-3992.1989.tb00315.x>
- Sürücü, L., & Maslakçi, A., (2020). Validity and reliability in quantitative research. *Business & Management Studies: An International Journal*, 8(3), 2694–2726. <https://doi.org/10.15295/bmij.v8i3.1540>
- Taie, S., & Goldring, R. (2020). *Characteristics of public and private elementary and secondary school teachers in the United States: Results from the 2017–18 national teacher and principal survey first look*. U.S. Department of Education. <https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2020142>
- Tennessee Department of Education (n.d.-a). *Graduation Requirements*. <https://www.tn.gov/education/families/graduation-requirements.html>

Tennessee Department of Education (n.d.-b) *Tennessee HOPE scholarship*.

<https://www.tn.gov/collegepays/money-for-college/tn-education-lottery-programs/tennessee-hope-scholarship.html>

Tennessee Department of Education. (n.d.-c). *Overview of Testing in Tennessee*.

<https://www.tn.gov/education/districts/lea-operations/assessment/testing-overview.html>

Tennessee Department of Education. (2022). *District Accountability*.

<https://www.tn.gov/education/districts/lea-operations/accountability/2022-district-accountability.html#:~:text=Under%20Tennessee's%20accountability%20system%2C%20districts,indicators%20measured%20through%20multiple%20pathways.>

Tennessee Education Glossary (n.d.). *Economically-disadvantaged students*. Tennessee

Comptroller of the Treasury. <https://comptroller.tn.gov/office-functions/research-and-education-accountability/collections/glossary.html>

Tennessee Education Association. (n.d.). *Cube Root Quick Score Formula and Calculation*

Procedures.

https://tnea.org/data/cms_files/old_website/sites/default/files/u98351/cube_root_quick_score_calculation.pdf

Tennessee S. Rep. No. 388. (2022).

<https://wapp.capitol.tn.gov/apps/BillInfo/Default.aspx?BillNumber=SB0388>

Tennessee School Board Association. (2016). *Board Policy and Operations*. [https://tsba.net/wp-](https://tsba.net/wp-content/uploads/2018/01/2016-17_BoardPolicy_Notebook.pdf)

[content/uploads/2018/01/2016-17_BoardPolicy_Notebook.pdf](https://tsba.net/wp-content/uploads/2018/01/2016-17_BoardPolicy_Notebook.pdf)

Tenny, S., Brannan, G.D., Brannan, J.M., & Sharts-Hopko, N.C. (2022, May 24). Qualitative Study. In *StatPearls*. StatPearls Publishing.

<https://www.ncbi.nlm.nih.gov/books/NBK470395/>

The Elements of Mental Tests. (2018, March 2). Classical test theory Park 1 (VE) Introduction [Video]. YouTube. <https://youtu.be/TfemL2xu1SU>

The Glossary of Education Reform. (2015, November 10). *Assessment*. Retrieved February 17, 2023 from <https://www.edglossary.org/assessment/>

Timmerman, E. (2012, January). *The Kalamazoo School Case*. Kalamazoo Public Library.

<https://www.kpl.gov/local-history/kalamazoo-history/kps/kalamazoo-school-case/>

Tomlinson, C. (2005). Grading and differentiation: Paradox or good practice? *Theory into Practice*, 44, 262-269. https://doi.org/10.1207/s15430421tip4403_11

Torpey, E.M. (2011). *You're a what? Psychometrician*. Bureau of Labor Statistics.

<https://www.bls.gov/careeroutlook/2011/fall/yawhat.pdf>

Tyner, A., & Gershenson, S. (2020). Conceptualizing grade inflation. *Economics of Education Review*, 78(2020). <https://doi.org/10.1016/j.econedurev.2020.102037>

U.S. Department of Education. (n.d.). Every student succeeds act (ESSA).

<https://www.ed.gov/essa?src=rn>

University of California San Francisco Office of Diversity and Outreach. (2023). *Unconscious Bias Training*. <https://diversity.ucsf.edu/programs-resources/training/unconscious-bias-training>

University of South Carolina (n.d.-a). *Before you begin grading*. Center for Teaching Excellence.

https://sc.edu/about/offices_and_divisions/cte/teaching_resources/grading_assessment_to_olbox/before_begin_grading/index.php

University of South Carolina (n.d.-b). *What is a grade?* Center for Teaching Excellence.

https://sc.edu/about/offices_and_divisions/cte/teaching_resources/grading_assessment_to_olbox/what_is_grade/index.php

Vatterott, C. (2015). *Rethinking grading*. ASCD.

Westphal, A., Vock, M., & Kretschmann, J. (2021). Unraveling the relationship between teacher-assigned grades, student personality, and standardized test scores. *Frontiers in Psychology*, 12(March). <https://doi.org/10.3389/fpsyg.2021.627440>

Witte, R.S., & Witte, J.S. (2017). *Statistics* (11th ed.). Wiley.

Wormeli, R. (2006). *Fair isn't always equal: Assessing and grading in the differentiated classroom*. Stenhouse Publishers.

Wormeli, R. (2012). It's time to stop averaging grades. *Middle Ground*, 16(2), 40-41.

<https://www.proquest.com/scholarly-journals/time-stop-averaging-grades/docview/1282883115/se-2>

Wraga, W.G. (2000, April 24-28). *The comprehensive high school in the United States: A historical perspective* [Paper presentation]. American Educational Research Association Annual Meeting, New Orleans, LA, United States.

Wright, B.D. (1977). Solving measurement problems with the rasch model. *Journal of Educational Measurement*, 14(2), 97-116.

Yale Poorvu Center for Teaching and Learning. (2021). *Formative and Summative Assessments*.

<https://poorvucenter.yale.edu/Formative-Summative-Assessments>

Yeager, D.S., Hanselman, P., Walton, G.M., Murray, J.S., Crosnoe, R., Muller, C., Tipton, E., Schneider, B., Hulleman, C.S., Hinojosa, C.P., Paunesku, D., Romero, C., Flint, K., Roberts, A., Trott, J., Iachan, R., Buontempo, J., Yang, S.M., Carvalho, C.M., Hahn, P.R., Gopalan, M., Mhatre, P., Ferguson, R., Duckworth, A.L., Dweck, C.S. (2019, September 19). A national experiment reveals where a growth mindset improves achievement. *Nature*, 573, 365-380. <https://doi.org/10.1038/s41586-019-1466-y>

Yeh, S.S. (2019). Origin of the achievement gap: Evidence from a league table analysis of 23 interventions. *Journal of Education Finance*, 45(1), 23-41.

<https://muse.jhu.edu/article/747803>

Zoeckler, L.G. (2007). Moral aspects of grading: A study of high school English teachers' perceptions. *American Secondary Education*, 35(2), p. 83-102.

APPENDIX: INTERVIEW QUESTIONS

Interviewer will open the interview with formal introductions and reminding the interviewee of the purpose of the interview. The interviewee will be asked to share their name, how many years they have been teaching, and what content area they teach. The interviewer will ask questions in italics.

Introduction: I am researching teacher experiences when they select a grading policy for their classroom as part of my research for my doctoral dissertation. This is part to fulfill part of the requirements of my program of study in ETSU's doctoral program for Educational Leadership. I will be recording this interview so that I can transcribe it for data analysis. Once the transcription is complete and pseudonyms are assigned, the recordings will be deleted. If you want to stop this interview at any time or would like me to stop recording at any point, please let me know. Do you have any questions?

Please share your name, where you teach, and what you teach.

Research Question 1: How do teachers perceive the use of grades in secondary English and math classrooms?

- 1. Please describe the purpose of grades in a high school classroom.*
- 2. What factors should be included in a grade?*

Research Question 2: What experiences are associated with choosing a grading practice for secondary English and math classrooms?

- 1. Please describe any coursework you had during your preservice program that focused on grading practices or measurement.*
- 2. Please describe any support or collaboration you had with others around developing your own grading practice. How did this impact your decision-making?*

3. *Please describe your own experience when selecting your own grading practice. How did you determine your final practice?*

Research Question 3: What experiences have teachers had with professional development centered on grading practices?

1. *What learning opportunities have you had for grading since you've been in the classroom?*
2. *Please describe any support or collaboration you had with others around developing your own grading practice.*

Research Question 4: What external influences affect teacher choice in grading practices?

1. *Please describe the grading culture in your school and/or district.*
2. *What external factors impact grading in your department or school and what does this impact look like?*

VITA

TALLYE WEHENKEL GASS

Education: Ed.D. Educational Leadership: East Tennessee State University;
Johnson City, TN; 2023
Ed.S. Educational Leadership: Lincoln Memorial University;
Harrogate, TN; 2010
M.A.T., Elementary Education (K-8): East Tennessee State
University; Johnson City, Tennessee; 2002
B.A. Communications: East Tennessee State University; Johnson
City, Tennessee; 2015
Public Schools: Greeneville, Tennessee

Professional Experience: Instructional Specialist, Greeneville City Schools: Greeneville,
Tennessee; 2019-Present
Teacher, Greeneville Middle School: Greeneville, Tennessee;
2011-2019
Teacher, East Ridge Middle School: Morristown, Tennessee;
2006-2011
Teacher, Greene County Schools: Greeneville, Tennessee; 2004-
2006
Graduate Assistant, East Tennessee State University, College of
Education; 2002

Honors and Awards: Robert E. Jordan Character Education Award