Digital Commons@Lindenwood University

Dissertations   Theses & Dissertations

Spring 4-25-2019

An Investigation of Patterns Between Demographic Characteristics of Teachers and Grading Practices

Jenelle Lee
Lindenwood University

An Investigation of Patterns Between Demographic Characteristics

of Teachers and Grading Practices

Lindenwood University Ed.D.

by

Jenelle Lee

A Dissertation submitted to the Education Faculty of Lindenwood University

in partial fulfillment of the requirements for the

degree of

Doctor of Education

School of Education

An Investigation of Patterns Between Demographic Characteristics

of Teachers and Grading Practices

Lindenwood University Ed.D.

by
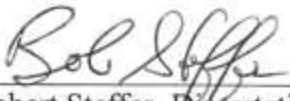
Jenelle Lee

This dissertation has been approved in partial fulfillment of the requirements for the

degree of

Doctor of Education

at Lindenwood University by the School of Education

<table>
<tr><td>_____</td><td>4-25-19</td></tr>
<tr><td>Dr. Robert Steffes, Dissertation Chair</td><td>Date</td></tr>
<tr><td>_____</td><td>4.25.19</td></tr>
<tr><td>Dr. John Long, Committee Member</td><td>Date</td></tr>
<tr><td>_____</td><td>4-25-19</td></tr>
<tr><td>Dr. Roger "Mitch" Nasser, Committee Member</td><td>Date</td></tr>
<tr><td>_____</td><td>4-25-19</td></tr>
<tr><td>Dr. Sherrie Wisdom, Committee Member</td><td>Date</td></tr>
</table>

Declaration of Originality

I do hereby declare and attest to the fact that this is an original study based solely upon my own scholarly work here at Lindenwood University and that I have not submitted it for any other college or university course or degree here or elsewhere.

Full Legal Name: Jenelle Ann Lee

Signature: _____ Date: 4/25/2019

Abstract

This mixed-methods phenomenological study examined a sample of grade 7 through 12 core content area teachers at four Missouri school districts. This study investigated the patterns linking demographic characteristics of teachers and grading practices. These practices were cross-referenced to the level of implementation of best-grading practices as it pertained to the amount of time and support systems a district provided for its teachers. The teachers in this study responded to a survey and provided insight regarding their grading practices. The purpose of the survey was to collect demographic information on the teachers, information on the academic and non-academic factors included in their grading practices, and information pertaining to the amount of time and support systems their districts provided to facilitate implementation of best-grading practices. Interviews with the teachers provided further insight into their grading practices. This study was derived from the beliefs that teachers can serve as agents of change in schools and that sustainable reform in grading practices benefits all stakeholders within the educational community. This study looked at grading practices from the teachers' perspectives to provide possible direction and support for schools that sought to implement best-grading practices and ensure fair and accurate grading.

Table of Contents

List of Tables

Chapter One: Introduction

Background of the Study

Teachers had a long history of using letter grades as the primary method by which to communicate their assessment of student mastery, dating back to the early 1900s (Moll, 1998). Although the terms associated with the topic of grading were frequently used interchangeably, they had distinct meanings (McTighe & Ferrera, 1998). Marks and scores were synonymous; they referred to the number score or letter given to any student test or performance. Grades (either in number or letter form) constituted what was reported at the end of a period of time as a summary statement of student performance (O'Connor, 2012).

Teachers also had a long history of using varying grading practices to calculate grades. Early 20th-century research of teachers' grading practices found that the reliability of the marks (later known as scores) calculated in the grades that were entered on report cards for the same material could be marked differently, depending on which teacher completed the marking (Cizek, Fitzgerald, & Rachor, 1996; Moll, 1998; Starch & Elliott, 1912, 1913a, 1913b). Research conducted near the start of the 21st century found that grading practices remained highly inconsistent (Allen, 2005; Cizek et al., 1996; Guskey, 2006).

A historically consistent, yet largely unaddressed problem resulting from teachers' use of varying grading practices was that the governing meaning of a student's grade was made ambiguous to interested parties, such as teachers, parents, colleges, employers, and even to the students themselves (Allen, 2005; Cizek et al., 1996). This made it challenging to possess a high degree of confidence that the factors included in a

teacher's grading practices resulted in accurate measurement of student mastery. As a result, the grades and marks teachers assigned to students had long been regarded by measurement experts as unreliable (Brookhart, 1993; Guskey, 2009; Stiggins, Frisbie, & Griswold, 1989). Clymer and Wiliam (2007) stated, "Current grading practices don't do the one thing they are meant to do, which is to provide an accurate indication of student achievement" (p. 36).

Few topics within education remained as controverted as the practice of grading (Leyrer, 2015). There had been much debate between teachers and between teachers and experts about the specific factors that should have been included when calculating grades. Academic factors were those considered to relate to a student's learning achievement (Guskey, 2002); whereas, non-academic factors were those that were included by teachers in grade calculation that did not directly relate to a student's achievement, but may have played an influential role in grade determination. A synthesis of research showed that examples included: organization, ability, aesthetic appearance of work, attendance, behavior, difficulty level of an assignment, effort, attitude, motivation, extra credit, homework, completion, improvement, participation, punctuality, and responsibility (Bailey, 2012; Brookhart, 1994, 2009; Guskey, 2004; McMillan, 2001, 2002; O'Connor, 2007).

Researchers found there was a level of personal ownership inherent in a teachers' grading practices that was deeply rooted in their personal and professional history. But, little was known about how teachers who used best-grading practices arrived at the point of adopting such practices and implementing them with fidelity (Leyrer, 2015). In addition, researchers found that, although some teachers were aware of the grading

practices of their colleagues next door (Allen, 2005), few teachers had knowledge about

the effectiveness of the grading practices employed by their colleagues (Leyrer, 2015).

Despite the ambiguities surrounding the use of varying grading practices, teachers

continued to possess a considerable amount of freedom in exercising personal judgment

about what factors they included when grading (Brookhart, 1994).

One reason grades were difficult to interpret was teachers exercised judgment

when they assigned grades and considered additional factors, other than achievement,

when they graded students (Brookhart, 1993; Zoeckler, 2007). Brookhart (1991) found

that teachers considered a "hodgepodge" of factors, consisting of both academic and non-

academic factors, when they calculated a student's grade (p. 36). The factors that each

teacher considered when grading depended on the individual teacher. Teachers usually

drew from their personal experiences as students that they regarded as being fair and

reasonable in determining the grading practices they employed (Bailey, 2012; Frary,

Cross, & Weber, 1992; Guskey, 2006, 2009; Guskey & Bailey, 2001; Leyrer, 2015).

Because recollections of these experiences varied among teachers, the grading practices

they employed varied as well (Guskey, 2006).

Since the inception of assigning letter grades, the act of grading was a significant

responsibility for educational professionals (Bailey, 2012; Chiekem, 2015; Guskey,

2004). Teachers' employment of varying grading practices was not only problematic to

the integrity of educational professionals, but to students as well. O'Connor (2002)

stated, "Grading that is faulty damages students and teachers" (p. 17). Teachers' grading

practices resulted in the assignment of grades that affected students' opportunities and,

ultimately, their futures (Leyrer, 2015). Several institutions used a student's grades in

making important educational, financial, and career decisions (Stiggins, 2001). Given the considerable impact grades had on a student's future, it was unsettling that few teachers received formal training in grading practices and the effectiveness of various grade calculation and reporting methods (Allen, 2005; Bailey, 2012; Brookhart, 2004; Guskey, 2004; Stiggins, 1993).

Purpose of the Dissertation

This mixed-methods phenomenological study examined a sample of grade 7 through 12 core content area teachers at four Missouri school districts. The main purpose of this study was to investigate possible patterns linking demographic characteristics of teachers and grading practices. The researcher cross-referenced these practices to the level of implementation of best-grading practices as it pertained to the amount of time a district provided for teacher collaboration, which included professional development, conversation with administration, and collaboration with peers. Specific demographics explored included: gender, age, content area, years of service, and educational level. This study investigated the relationships between teacher demographic characteristics, grading, and teacher collaboration time provided by the district.

The data from this study expanded upon existing research, and the researcher identified the factors used in grading by teachers possessing particular demographic characteristics to help ensure they were rating student performance in an appropriate and fair way (Carlson, 2003; Shanahan, 2011). The desire to improve teachers' grading practices to ensure fair and accurate grading for students was the motivation behind this study. In addition, this study expanded upon existing research to examine the factors that contributed to grades in an attempt to help close the then-current gap in understanding of

grading practices. This helped districts plan for professional development and move towards sustainable school reform.

Several educational reform leaders stated that classroom teachers did not follow many of the recommended practices for grading (Barnes, 1985; Brookhart, 1993; Manke & Loyd, 1990; Shanahan, 2011; Stiggins & Conklin, 1992). Instead, researchers found that teachers used judgment when assigning grades (Brookhart, 1993) and considered non-academic factors against expert recommendation (Brookhart, 1994; Bursuck et al., 1996; Cross & Frary, 1996; McMillan, Myran, & Workman, 2002; Shanahan, 2011; Strein & Meshbesher, 2006).

Research further indicated that when researchers investigated teacher grading practices, they generally asked teachers to measure the degree to which they incorporated different factors into their grading (Bursuck et al, 1996; Cross & Frary, 1996; Frary et al., 1992; McMillan et al., 2002, Shanahan, 2011). However, teachers were not consistently asked to indicate their use of non-academic factors when assigning grades (Bursuck et al., 1996; Shanahan, 2011). Some studies concluded that non-academic factors accounted for some degree of the variance in grades (Lekholm & Cliffordson, 2008; Shanahan, 2011) and other studies concluded that many non-academic factors were related to the grades students received (Bruckman, 2010; Duckworth & Seligman, 2006; Gottfredson, 1981; Hinshaw, 1992; Randall & Engerhard, 2010; Shanahan, 2011).

This study expanded further upon the demographic characteristics of teachers choosing to include specific non-academic factors in grades. The literature supported that discrepancies existed between teachers' grading practices and measurement experts' recommendations (Allen, 2005; Bailey, 2012; Brookhart, 1994; Stiggins, Frisbie, &

Griswold, 1989). However, research examining the extent to which teachers' grading

practices reflected the constructs they were instructed to incorporate into their grading

was scarce (Shanahan, 2011). Therefore, there was a then-current gap in the

understanding of grading practices and grading policies; thereby, establishing a need to

further investigate the academic and non-academic factors that contributed to grades

(Bailey, 2012; Shanahan, 2011).

Rationale

   While research recognized an emphasis on academic achievement by teachers in

grading, it also recognized that academic achievement was not the only factor considered

by teachers in grading (Brookhart, 1994; Bursuck et al., 1996; Cross & Frary, 1996;

McMillan et al., 2002; Shanahan, 2011; Strein & Meshbesher, 2006; Zoeckler, 2007).

Many teachers combined academic factors with non-academic factors to create a single

letter grade, because they felt it was necessary to provide stakeholders with a more

complete representation of student performance than was provided by simply reporting

student academic achievement (Guskey, 1996). Researchers found that teachers often

included non-academic factors in grades as a means of communicating various messages

about "level of expectation, level of academic achievement, encouragement, and

disappointment" (Zoeckler, 2007, p. 97).

   However, when teachers mixed academic and non-academic factors into a single

letter grade, it quashed the grade's ability to clearly communicate any one aspect of a

student's education (Guskey, 2001; Leyrer, 2015; Marzano, 2010; Stiggins, 2001). This

resulted in grades that were inaccurate measures of student mastery after they had been

contaminated with non-academic factors, such as attendance, effort, behavior, and ability

(Allen, 2005; Greville, 2009; Shanahan, 2011). Their inclusion in grading allowed greater

potential for discrepancies between teachers' grades, due to subjective teacher bias,

thereby having distorted the meaning of a grade (Bailey, 2012; Cross & Frary, 1999). As

a result, experts recommended that teachers reported information pertaining to non-

academic factors separate from academic factors (Brookhart, 2004; Guskey & Bailey,

2001; Marzano, 2010; O'Connor, 2009; Scriffiny, 2008; Wiggins & McTighe, 2006;

Winger, 2005).

Some districts adopted grading policies in an attempt to address inconsistent

grading practices (Polloway et al., 1994; Reeves, 2008; Strein & Meshbesher, 2006;

Shanahan, 2011). While effective grading practices were proposed, they were not adopted

on a widespread scale (Cross & Frary, 1999; O'Connor, 2007). Furthermore, researchers

found that even in districts possessing policies to guide teachers' grading practices, many

teachers ignored the policies and included factors outside of the criteria established

within them (Bruckman, 2010; Buzzelli & Johnston, 2002; Shanahan, 2011; Strein &

Meshbesher, 2006). However, despite personal and professional histories and long

traditions of varying grading practices used by teachers, educators could change their

grading practice guidelines and policies to more effective ones, under certain influences

(Dyd, 2012). This change entailed both a shift away from imprecise systems that

combined academic and non-academic factors into a single letter grade, and a move

toward grading practices designed to measure academic achievement (academic factors)

separately from non-academic factors ("Effective Grading Practices," 2011) pertaining to

behavior, work habits, and attitude (Brookhart, 2009).

This empirical study provided administrators with a resource to explore how

teachers implemented best practices in grading. Since schools used data (formative and

summative) to drive change, it was vital that the data were an accurate reflection of

student knowledge. Inaccurate grading practices undermined the ability of administrators

to address the professional development needs of staff members (Dyd, 2012; Wilson,

2004). This study informed conversations about grading practices and the level of

implementation of best practices, as it pertained to the amount of time a district provided

for teacher collaboration. By conducting an analysis on the amount of time districts

provided for teacher collaboration (which included professional development,

conversation with administration, and collaboration with peers) and comparing it to

teacher grading practices, the researcher hoped to bring awareness to these two issues.

Given the results of this study, administrators and educational professionals might have

used the information to ensure fair and accurate grading.

Research identified 15 best practices for grading and reporting. These practices

were referred to by O'Connor (2007) as the 15 fixes for broken grades. They included the

following: (a) Do not include student behaviors in grades- include only achievement; (b)

Do not reduce marks on late work- provide support for the learner; (c) Do not give extra

credit or bonus points; (d) Do not punish academic dishonesty with reduced grades- apply

other consequences; (e) Do not consider attendance in grade determination- report

absences separately; (f) Do not include group work in grades; (g) Organize and report

evidence by standards/learning goals; (h) Provide clear descriptions of achievement

expectations; (i) Compare each student's individual performance to preset standards; (j)

Rely on quality assessments; (k) Do not rely only on the mean for central tendency- use

professional judgment; (l) Do not punitively assign zeroes for missing work- use

alternatives; (m) Use only summative assessment evidence to determine grades; (n) Do not summarize evidence accumulated over time- emphasize more recent achievement; and (o) Students should play key roles in assessment and grading to promote achievement. Throughout the study the researcher used these practices as a model for best-grading practices and as a guide to investigate then-current realities.

Links existed between teacher perceptions and teacher actions that validated the importance of understanding teacher perceptions of grading (Bailey, 2012; Hardre' et al., 2006). Research suggested that teacher perceptions of their students impacted the approach teachers utilized with their students (Bailey, 2012; Biddle & Anderson, 1986; Wenglinski, 2000). In addition, research suggested that what teachers did influenced student academic achievement (Greene, Miller, Crowson, Duke, & Akey, 2005). Synthesis of the impacts of student-teacher relationships supported the need to further expand research in these areas.

Research Questions and Hypotheses

H1: There will be differences in demographic characteristics within respondents to a Likert-scale survey, with regard to best-grading practices, among the following teacher-related variables: Teachers' age, teachers' gender, teachers' levels of education, teacher's years of service, and teachers' content areas.

H2: There will be differences in perception of best-grading practices, measured by a Likert-scale survey, among the following support structures: Frequency of teachers' discussions with peers regarding best-grading practices, frequency of teachers' discussions with building administrators regarding best practices, and frequency of the teachers' engagement in self-reflections in which the teacher is engaged.

H3: There will be differences in perception of agreement topic-to-topic, and agreement-to-disagreement with survey statements, with regard to best-grading practices, measured by a Likert-scale, with regard to the following support characteristics: A student's behavior, grades as a means of punishment, student responsibility, a student's ability level, assignment completion points, extra credit, a teacher's personal judgment, assignments other than assessments, a student's involvement in extracurricular activities, zeroes for missing assignments, a student's GPA, class participation, a student's gender, a student's physical appearance, homework, note-taking, aesthetic appearance of a student's work, a student's personality, student effort, using grades as a reward, student progress towards achieving the course learning standards, individual grades for group work, individual projects, project-based grading, project-based assignments, student portfolios, conferences with students about individual grades throughout each grading period, student-set personal learning goals for the course, student personal reflection on achievement of the course learning standards, district-provided time for teachers in the same content area to meet to align grading practices, district-provided professional development about best-grading practices, pre-service training for teachers on best-grading practices, the belief that teachers should be accountable for following best-grading practices, teacher-to-teacher grading practice comparisons, district-provided common planning time for teachers in the same content area, district-provided collaboration time for teachers on early release days, district-provided collaboration time for teachers on late start days, district-provided collaboration time for teachers on professional development days, teacher collaboration, the belief that grades should be a direct reflection of student mastery of course learning standards, grades as a means to

provide feedback to stakeholders, gradebook field that indicates student mastery of

learning competencies, re-do assignments, test re-takes, and summative and formative

assessments.

H4: There will be relationships among levels of agreement topic-to-topic, with

regard to best-grading practices, measured by a Likert-scale survey, with regard to the

following support characteristics: A student's behavior, grades as a means of punishment,

student responsibility, a student's ability level, assignment completion points, extra

credit, a teacher's personal judgment, assignments other than assessments, a student's

involvement in extracurricular activities, zeroes for missing assignments, a student's

GPA, class participation, a student's gender, a student's physical appearance, homework,

note-taking, the aesthetic appearance of a student's work, a student's personality, student

effort, using grades as a reward, student progress towards achieving the course learning

standards, individual grades for group work, individual projects, project-based grading,

project-based assignments, student portfolios, conferences with students about individual

grades throughout each grading period, student-set personal learning goals for the course,

student personal reflection on achievement of the course learning standards, district-

provided time for teachers in the same content area to meet to align grading practices,

district-provided professional development about best-grading practices, pre-service

training for teachers on best-grading practices, the belief that teachers should be

accountable for following best-grading practices, teacher-to-teacher grading practice

comparisons, district-provided common planning time for teachers in the same content

area, district-provided collaboration time for teachers on early release days, district-

provided collaboration time for teachers on late start days, district-provided collaboration

time for teachers on professional development days, teacher collaboration, the belief that grades should be a direct reflection of student mastery of course learning standards, grades as a means to provide feedback to stakeholders, gradebook field that indicates student mastery of learning competencies, re-do assignments, test re-takes, and summative and formative assessments.

RQ1: What types of support work best for teachers trying to implement best-grading practices?

RQ2: Which academic factors do teachers use when grading?

RQ3: Why do teachers choose to use academic factors when grading?

RQ4: Which non-academic factors do teachers use when grading?

RQ5: Why do teachers choose to use non-academic factors when grading?

Limitations

The outcome of this study may be affected by the following limitations:

1) The study was limited to junior/high schools in one Midwestern state. Due to this limited population, the results may not be generalizable nationwide.

2) The study included dependent variables that were self-reported data. The data were contingent upon the personal perceptions of each teacher. If teachers viewed that their perceptions of grading practices were deemed as positive or negative in the context of the study, they may have altered their responses.

3) The survey used in this study was designed to measure the perceptions of grades 7 through 12 core content area teachers who taught in grades 7 through 12. Therefore, the results may not be generalizable to grades K through 6. In this study, superintendents were asked to suggest the names of teachers to interview. The

superintendents may have acted with bias by selecting teachers whom they felt would have positively reflected their beliefs about the appropriate relationship between grading practices and the independent variables studied.

Assumptions

1) It was an assumption that measuring teacher perceptions through self-reporting was an accurate method of predicting actual teacher behaviors regarding demographic characteristics and grading practices.

2) It was an assumption that participants in this study were certified to teach in the content area and grade level that they reported.

3) It was an assumption that participants in this study were not deceptive with their answers, and that the participants answered questions honestly and to the best of their ability.

Definition of Terms

  Academic Achievement. A student's mastery of specific learning standards (Wormeli, 2006).

  Academic Factors. Those that were included by teachers in grade calculation student achievement that represented mastery of academic content (Wormeli, 2006).

  Best-Grading Practices. This referred to the 15 best practices also known as the "Fifteen Fixes for Broken Grades" by O'Connor (2007).

  Collaboration. The time a district provided for professional development, conversation with administration, and collaboration with peers.

Grade. Constituted what was reported at the end of a period of time as a summary statement of student performance- presented in either number or letter form (O'Connor, 2012).

Grading Practices. The methods teachers used to determine student grades, which included the factors contributing to the formation of grades (Bailey, 2012; Brookhart, 2009).

Non-Academic Factors. Those that were included by teachers in grade calculation or that influenced a teacher's assignment of a student's grade. A synthesis of research included the following examples: ability, aesthetic appearance of work, attendance, attitude, behavior, completion, difficulty level of an assignment, effort, extra credit, improvement/progress made, motivation, organization, homework, participation, punctuality, responsibility, and student personality (Brookhart, 1994, 2009; Cross & Frary, 1999; Erickson, 2010; Guskey, 2009; Guskey & Bailey, 2001; O'Connor, 2007; McMillan, 2001, 2002).

Non-Veteran Teacher. A teacher who has been working less than 10 years in a full-time teaching capacity in a public school setting (Feistritzer, 2011).

Professional Development. A comprehensive, sustained and intensive approach to professional learning that facilitates the collective responsibility of teachers and principals for improved student achievement (National Staff Development Council, 2009).

Veteran Teacher. A teacher who has been working 10 or more years in a full-time teaching capacity in a public school setting (Feistritzer, 2011).

Conclusion

Best practices for grading emerged due to the increasing demands of school improvement because of the need for accurate measurement of student performance. This empirical study investigated possible patterns between demographic characteristics of teachers and grading practices. The researcher cross-referenced these practices to the level of implementation of best-grading practices as it pertained to the amount of time a district provided for teacher collaboration which included professional development, conversation with administration, and collaboration with peers. The researcher felt confident that this study would have a positive contribution to school reform by expanding upon the 15 best-grading practices by O'Connor (2007).

Data were collected via teacher surveys. This was a mixed-methods, phenomenological study. The researcher used an electronic survey generator to collect and analyze the data. The survey and Letter to Participate were emailed to the superintendent of each of the schools in the 16 counties that were members of Missouri's Southeast Regional Professional Development Center (SE RPDC). Each superintendent was asked to forward the email to the seventh through twelfth-grade core content area teachers in their respective district. The goal was to have at least three teachers from each of the school districts complete the survey.

SE RPDC was comprised of schools in the following counties: Bollinger; Butler; Cape Girardeau; Carter; Dunklin; Madison; Mississippi; New Madrid; Pemiscot; Perry; Ripley; Scott; Ste. Genevieve; St. Francois; Stoddard; and Wayne. SE RPDC's mission was to improve student achievement for all students in Southeast Missouri by providing appropriate and timely professional development to school communities in their service

area (Southeast Missouri State University, 2019). Its member schools were focusing on a

Collaborative Work Project (CW Project), wherein they emphasized common formative

assessment, data-based decision making, and effective teaching/learning practices.

Chapter Two is a review of literature investigating pedagogical beliefs and

practices surrounding the realities of 21st century grading.

Chapter Two: The Literature Review

Introduction

For over a century, grades were a central component of American schools. While most students received grades from their teachers, most students were uninformed as to the specific factors considered by each of their teachers when calculating their grades. Despite their consistent presence in the educational system since the early 1900s, grades had not always looked the same, served the same purpose, nor created the same impact on students' futures (Schneider & Hutt, 2013).

Although society created "a huge edifice based on the grades and marks given by teachers . . . there is serious doubt about the validity and the reliability of those marks" (Fleming, 1999, p. 83). Marzano (2000) stated that there was considerable speculation regarding the value of grades as a result of teachers' inaccurate systems of varying grading practices concerning grade calculation and reporting. Measurement experts made grading practice recommendations for teachers that were largely ignored (Griswold, 1993). These recommendations were intended to support learning (Tomlinson, 2003) by separating academic and non-academic factors in grades (Guskey, 2004). In short, they were intended to eliminate inaccurate grading practices that interfered with the measurement, reporting, and feedback of student achievement (O'Connor, 2009).

Brookhart (1993) purported that, although most teachers understood that combining academic and non-academic factors in grading was an unacceptable practice, they continued with this method for several reasons. These reasons ranged from teachers possessing inadequate knowledge regarding best practices, due to a lack of training in their pre-service programs, to their tendency to replicate the practices they experienced

when they were students (Brookhart & Nikito, 2008; Guskey, 2009; Leyrer, 2015; Stiggins, 2002, 2008). Therefore, "Who teachers are, where they are coming from, and what they think is of great interest to every segment of society;" (Feistritzer, 2011, p. viii).

A review of the history and literature in terms of grading practices and the demographic characteristics of teachers is the focus of this chapter.

History of Grading

Marzano (2000) accredited Yale University and Mount Holyoke College as having established the origin of our current grading system, at the time of this writing. He stated that, in 1783, Yale began the shift from providing narrative feedback to using a four-point grading scale, by which they had assessed their students. This grading scale consisted of the Latin terms "optime," "second optime," "inferiors," and "pejores" (Webster, 2011, p. 17). By 1877, Mount Holyoke changed these divisions to 'A,' 'B,' 'C,' 'D,' and 'F,' symbols that became commonly used by universities (Eison, Pollio, & Milton, 1986). By the early 1920s, the 'A-F' grading system had first spread into the high schools and would later spread to the majority of schools at every level (Eison et al., 1986).

Prior to the 1900s, students outside of post-secondary educational institutions were not given grades. Instead, teachers provided students with meaningful feedback on their performance through narrative comments (Marzano, 2000). Around the early 1900s, the American population shifted from rural to large urban centers that challenged the efficiency of the one-room schoolhouse, wherein students of all ages and backgrounds were grouped together with one teacher (Moll, 1998). The implementation of mass

compulsory education laws in the United States during the late 19th and early 20th centuries strengthened the connection between the schoolhouse, great universities of the east, and the labor market (Schneider & Hutt, 2013). Along with the population shift and increased enrollment came the introduction of the uniform report card with marks (later referred to as grades) that indicated a teacher's assessment of a student's performance to interested parties.

Initially, teachers used grades to communicate internally with families and other teachers within the same school. But, after the change to mass compulsory schooling and an increasingly industrial economy, teachers began using grades to communicate externally with college admissions departments and businesses for the purpose of distinguishing the aptitude levels between students and potential workers (Schneider & Hutt, 2013). In addition, grades also helped track students who moved from one school to another as population migration increased throughout this time (Snyder, 1993). In 1870, there were 500 reported high schools, and by 1910 this number had grown to 1,900, resulting in an exponential increase in the number of students, as well (Guskey & Bailey, 2001).

As grades grew in importance, so too did researchers' interests in grades. Researchers struggled with the purpose and practices of grading for more than a century (Guskey & Bailey, 2001). Beginning in 1911, researchers testing the reliability of the marks entered on report cards concluded that the same material could be marked extremely different, depending on which teacher was doing the marking (Cizek et al., 1996; Moll, 1998; Starch & Elliott, 1912, 1913a, 1913b). In addition, there was no uniformity as to how often grades were issued (Ashbaugh & Chapman, 1925). By the

time these inconsistencies were discovered, the report card was already widely embraced

and was commonly used as the medium by which teachers reported student progress

(Moll, 1998).

By the 1930s, it became apparent to researchers that there were major problems

with both the grading system and teachers' grading practices when research, such as that

conducted by Rinsland (1937), concluded, "When all is said and all studies examined,

one is forced to admit that the whole grading system is highly subjective, unreliable and

unfair" (p. 26). The field of child psychology emerged at this time and raised additional

concern about the topic of grades and their potentially adverse effects on the concept of

the whole child and children's self-images (Webster, 2011). However, such findings had

virtually no impact. Throughout 1911 to 1960, schools still continued to use the report

card, but experimented with various letter and number reporting conventions, by which

they indicated a teacher's assessment of student achievement (Cizek et al., 1996; Moll,

1998).

In the latter half of the 19th century and early part of the 20th century, percentage

grading was the most commonly used system (Moll, 1998). In this system, teachers

assigned each student a number between zero and 100 that they felt represented the

percentage of material the student had learned. Marks below 50% were considered failing

because they were seldom assigned (Moll, 1998). During the 1930s and 1940s, most

schools switched from percentage grading to assigning letter grades consisting of 'A,'

'B,' 'C,' 'D,' and 'F,' that represented groups of percentages (Cizek et al., 1996; Moll,

1998). A percentage grade of 50 or less distinguished a student as failing (Guskey, 2013).

In the 1960s, some schools began using pass/fail options and written evaluations to make

their grades less bureaucratic and more humanized (Moll, 1998). However, most schools

at this time still continued to use what had become regarded as the traditional practice of

using the 'A-F' letter system, wherein letter grades were changed into numbers to

calculate a student's grade point average (GPA) (Burke, 1968; Chansky, 1964; Schneider

& Hutt, 2013). During the 1960s, the letter grade system took hold (Schneider & Hutt,

2013) and became a principal fixture in the educational system.

Concerns about the validity and consistency of grading practices continued to

remain and by 1964 grades still represented different things to different teachers

(Chansky, 1964). A 1968 study by Temple University researchers revealed that grading

practices were a problem at the collegiate level as well (Schneider & Hull, 2013). The

researchers found that different professors teaching the same course produced

substantially varying grades. Similar to that of previous research, this finding had

virtually no impact. Instead, schools became dependent upon a standardized grading

system to motivate students, to determine placement, and to communicate about student

learning both internally and externally (Schneider & Hull, 2013).

Some schools modified the letter grade system to include a plus or minus

(Herman, 2013). Other schools included symbols "(e.g., S = Satisfactory, N = Needs

Improvement, U = Unsatisfactory)" or descriptors "(e.g., Emerging, Developing,

Maturing)" in place of letter grades (Cizek et al., 1996, p. 161). Despite numerous

experiments and modifications in grading systems since the early 1900s, research showed

that letter grades remained the most common grading practice in schools (Moll, 1998).

Regardless of the kind of mark used by teachers to indicate student mastery, grades

continued to serve as a method by which to communicate important information to stakeholders about student performance and progress (Cizek et al., 1996).

Throughout America, students were expected to learn a diverse array of content, and teachers were expected to assess students' knowledge of this content and summarize the learning by assigning students a letter grade (Allen, 2005). Common practice was for teachers to record students' scores on individual assignments in a grade book and average the scores together at the end of the grading period to calculate a final grade (Erickson, 2010). While this might have appeared like a straightforward process, the nature of the individual assignments (academic and non-academic) included the scores for the individual assignments, and the final grades considerably varied depending on the individual teachers ("Effective Grading Practices," 2011).

Therefore, a historically consistent problem largely unaddressed since the inception of grades was that teachers' practices for assigning grades differed significantly and unpredictably (Cizek et al., 1996). Brookhart (1994) concluded that teachers' grading practices usually "confound constructs into composite scores of questionable reliability and validity and thus uncertain meaning" (p. 299). This resulted in an ambiguous "meaning of a student's grade to any interested party- the parents, college admissions departments, employers, and even the student" (Cizek et al., 1996, p. 161).

Despite their important communication role in the educational process for students ranging from kindergarten to university, grades continued to be a source of confusion and were commonly misunderstood by stakeholders (Allen, 2005; Shanahan, 2011). In addition, varying grading practices had caused some teachers to have a reputation as being "hard graders," versus others who were regarded as "easy graders"

(Zoeckler, 2007, p. 100). Stakeholders attributed the difference between "hard graders" versus "easy graders" to a teacher's personality and often viewed grading to be "a matter of arbitrary caprice" (Zoeckler, 2007, p. 100).

Ebel and Frisbie (1986) attributed much of the controversy surrounding teachers' grading practices to three factors: (a) The technical challenges of accurately measuring achievement; (b) variations in educational philosophies among teachers; and (c) the conflict in roles teachers faced when they acted as both advocates for, and judges of, their pupils. These factors made it challenging to possess a high degree of confidence that the factors included in a teacher's grading practices resulted in accurate measurement of student mastery. Reeves (2006) asserted, "The freedom long enjoyed by private practitioners must take place within boundaries of fairness, mathematical accuracy, and effectiveness" (pp. 113-114).

To further account for variation in teachers' grading practices, research conducted by Brookhart (2009) found there were three types of grading: (a) criterion-referenced; (b) self-referenced; and (c) norm-referenced. Criterion-referenced grading constituted comparing a student's work to standards, self-referenced grading constituted comparing a student's work to his/her previous work to measure progress. Whereas, norm-referenced grading constituted comparing a student's work to other students' work (Brookhart, 2009).

No matter the type of grading used, the validity of the grade was paramount (Allen, 2005). Allen (2005) wrote, "Validity addresses the accuracy of the assessment and grading procedures being used by teachers" (p. 218). The validity of a grade was essential because

the sole purpose of the grade was to accurately communicate to others the level of

academic achievement that a student has obtained; If the grades are not accurate

measures of the student's achievement, then they do not communicate the truth.

(Allen, 2005, p. 218)

While there had been substantial improvements in the range and quality of

information about educational performance available to interested parties, teachers'

grading practices remained as varied as they were in the early 1900s (Cizek et al., 1996).

Teachers continued to be challenged by the task of rating student performance in a

manner that was both equitable and appropriate (Carlson, 2003; Shanahan, 2011). One

reason why grades could have been difficult to interpret was that teachers exercised

judgment when assigning grades and considered additional factors, other than

achievement, when grading students (Brookhart, 1993; Marzano, 2000). Brookhart

(1991) found that teachers considered both academic and non-academic factors in their

grading practices, because they strongly believed that non-academic factors were relevant

to grading. However, researchers also found that teachers rarely specified all components

of a grade that they took into consideration (Guskey, 2009).

Experts considered academic factors as those that pertained to a student's learning

achievement (Guskey, 2002). Non-academic factors were those considered to have

pertained to a student's behavior, attitude, and effort. Non-academic factors were things

that could have potentially enhanced or diminished academic factors and/or achievement.

Teachers often combined academic and non-academic factors when grading in an attempt

to give stakeholders a more complete depiction of a student's level of achievement. But

experts cautioned against doing so because it blurred the intended meaning of the grade -

which experts stated should have only been based on academic factors (Guskey, 2001).

Due to their mixing of academic and non-academic factors in grading, teachers likely continued to struggle with the development of "meaningful, reasonable, and equitable grading practices and policies" (Guskey, 2006, p. 674). Guskey (2006) warned that this struggle would only compound in years to come, if teachers continued to use reporting systems wherein they combined diverse factors of evidence of student learning into a single grade. By separating academic and non-academic factors from a single grade, teachers could have made the meaning of grades less ambiguous and improved their communicative value (Guskey, 2006).

McMillan, Myran, and Workman (1999) stated another reason grades were difficult to interpret and caused their validity to be questioned was that several teachers used grades for diverse purposes (Leyrer, 2015). Guskey (2004) found that teachers routinely used grades to achieve the following purposes: (a) to communicate achievement to stakeholders; (b) to provide information for students as part of their self-evaluation; (c) to identify students for particular educational programs; (d) to incentivize students to learn; (e) to evaluate an instructional program's level of effectiveness; and (f) to give support of students' lack of effort and/or responsibility, or downturn in behavior expectations. Hendrickson and Gable (1997) found that teachers gave students positive grades (As and Bs) for four purposes: (a) To show a student's progress toward a predetermined goal; (b) to indicate that a student showed competence or achievement in a particular subject area; (c) to indicate that a student demonstrated effort toward achievement; and (d) to reflect that a student positively compared to other students of the same age.

No matter the purpose, many teachers thought both academic and non-academic factors needed to be included in grade calculation to accurately depict student performance (Guskey, 2004). As a result, Guskey (2006) purported that teachers should have used a "widely varying criteria to determine students' grades" (p. 672). These criteria consisted of both academic and non-academic factors that could have been classified into one of three categories of criteria: product (academic), process (non-academic), and progress (a student's individual gain deriving from their overall learning experience) (Brookhart, 1993; Frary, Cross, & Weber, 1993).

Product criteria emphasized the assessment of what students know and are able to achieve at a specific point in their academic career (Guskey, 2006). Teachers that used product criteria usually only included "final examination scores, final reports or projects, overall assessments, and other culminating demonstrations of learning" (Guskey, 2006, p. 672). Guskey (2008) stated, the act of teachers exclusively basing their grades on product criteria was typically the result of their compelling belief that communicating summative assessments of students' achievement was the most important purpose of grading. Therefore, grades based on product criteria alone were the preference of standards-based teaching and learning advocates (Guskey, 2006).

Conversely, process criteria were favored by teachers who disagreed with advocates of product criteria and their argument that it provided a more complete illustration of student learning (Guskey, 2006). Teachers using process criteria believed that "grades should reflect not only the final results but also how students got there" (Guskey, 2006, p. 672). Therefore, Guskey (2006) stated that teachers emphasizing process criteria in their grade calculations frequently included behavioral components,

such as effort or work habits in addition to formative assessments, such as "quizzes, homework, punctuality of assignments, class participation, or attendance" (p. 672).

Progress criteria was also referred to as "learning gain," "improvement scoring," "value-added learning," and "educational growth" (Guskey, 2006, p. 672). Teachers using progress criteria made distinctions between progress (which they determined by measuring backward from a final performance standard) and growth (which they determined by measuring forward from the student's initial position on a learning continuum) (Brookhart, 1999; Guskey, 2006). However, Guskey (2006) maintained, when achievement was assessed using "well-defined learning standards that includes graduated levels of performance, progress and growth criteria could be considered synonymous" (p. 272). As a result, teachers using progress criteria would consider how much improvement students achieved over an established timeframe, as opposed to where they were at one particular point (Guskey, 2006). Teachers using progress criteria were typically in fields of special education, because their grading criteria resulted in grades that may have been highly individualized (Guskey, 2006).

Researchers found that most teachers used some combination of the three types of criteria when determining a student's grade (Guskey, 2008). The National Center for Education Statistics (NCES) conducted a survey of 4,000 teachers in 1999 that showed this combination was common practice (Webster, 2011). The survey found that 76% of the teachers reported that they based grades on absolute achievement (product criteria/ academic factors); 97% based their grades on a student's effort (process criteria/non-academic factors); and 84% reported having used individual improvement (progress criteria) in calculating their grades.

Ornstein (1994) stated that several teachers varied their grading criteria based upon each individual student, making considerations as they saw fit. However, the practice of varying grades based on the teacher's personal judgment of particular students greatly compromised the meaning of any grade and had the potential to invite accusations of bias (Guskey, 2006). As stated by Rom and Musgrave (2014), "Grading bias is the result of incorporating illegitimate factors into an instructor's assessment of students' work" (p. 139). One repercussion of grading bias was that grades did not accurately represent the student's mastery of learning objectives (Rom & Musgrave, 2014). Another repercussion was that grading bias forced students to make a choice between either doing their best work or catering to their teacher's biases (Rom & Musgrave, 2014).

Grading bias could be conscious (Malouff, 2008) or unconscious (Malouff, Stein, Bothma, Coulter, & Emmerton, 2014) and presented in a variety of forms. Grading bias could also be presented as a result of a teacher's prior experience with a student (e.g. a halo or horn effect), as a result of a physical characteristic of a student (e.g. race, gender, physical attractiveness, etc.), or as a result of an assigned classification, such as gifted or learning impaired (Malouff & Thorsteinsson, 2016).

Researchers found that disagreement among teachers regarding the purpose of grades had a widespread impact throughout educational institutions and the educational system as a whole. Guskey (1996) stated that there were five purposes of grading: (a) To communicate the achievement status of students to parents, students, and others; (b) to provide information that students can use for self-evaluation; (c) to select, identify, or group students for certain educational paths or programs; (d) to provide incentives to learn and; (e) to evaluate the effectiveness of instructional programs. However, Brookhart

(2004) noted, "It is very difficult for one measure to serve different purposes well" (p.

21). Brookhart (2004) summarized the overall issue pertaining to the inclusion of

academic and non-academic factors in the grading process in stating, "The main

difficulty driving grading issues both historically and currently is that grades are pressed

to serve a variety of conflicting purposes" (p. 31).

Guskey (2006) believed varying grading practices were the result of ambiguity

among teachers about both the purpose of grading and the format used to report grades.

He stated that teachers struggled with decisions about which factors to include when

grading, largely because they were unclear as to the purpose their grades should really be

serving. Guskey (2006) also stated that, because most high school reporting forms only

allowed a single letter grade to be assigned to students enrolled in each particular course

or content area, teachers were often forced to combine diverse factors, consisting of both

academic and non-academic factors, into a single letter grade symbolizing student

mastery. It would not be until all of these issues were addressed, he warned, that grades

would become more meaningful.

Guskey (2006) stated that the first step in the process of developing fair and

accurate grading practices was for teachers to participate in candid discussions about the

purpose of grading and reporting. "Teachers should consider what message they want to

communicate through grading, who the primary audience for the message is and what the

intended goal of the communication is" (p. 52). Only after issues about purpose were

resolved could decisions about the appropriateness of particular grading policies be

resolved (Guskey, 2006).

O'Connor (2007) expanded upon Brookhart (2004) and Guskey (2006) by

contending that a shared vision of the purpose of grades among teachers within the same

school and entire school district was of the utmost importance in working to achieve

ideals, such as consistency and fairness. Marzano (2000) revealed that most teachers were

aware of inadequacies in their current grading and reporting methods. Yet, as Kain

(1996) concluded, teachers within the same school, let alone same district, failed to

communicate about grades, which inhibited collaboration and resulted in teachers

sticking to their individual grading practices. Therefore, it was of the utmost importance

that teachers communicated about the grading process to begin the process of addressing

such inadequacies. This process

> requires thoughtful and informed professional judgment, concern for what best
>
> serves the interests of students and their families, and careful examination of the
>
> tasks students are asked to complete and the questions they are asked to answer to
>
> demonstrate their learning. (Guskey, 2013, p. 72-73)

Researchers found there was a level of personal ownership inherent in a teacher's

grading practice that was deeply rooted in personal and professional history, with little

known about how teachers who used best-grading practices arrived at those specific

practices (Leyrer, 2015). In addition, although researchers found some teachers were

aware of the grading practices used by their next-door colleagues (Allen, 2005), few

teachers had knowledge about the effectiveness of such grading practices (Leyrer, 2015).

Despite the ambiguities and serious consequences created for students surrounding the

use of varying grading practices, teachers continued to possess a considerable amount of

freedom in exercising personal judgment about what factors they included when grading

(Brookhart, 1994). Wiggins and McTighe (2007) stated,

We believe teaching is currently far too personalized. Without long-term results

and shared analysis of goals to study together or shared standards of best practice

to which we refer, teachers have little choice but to (over)emphasize personal

beliefs, habits, and style. Naturally, then, any criticism of our teaching makes

most of us defensive and resistant to the message. (p. 111)

Grading reform initiatives were often hindered, due to many teachers' beliefs

about their ability to exercise personal judgment in their grading practices (Wiggins &

McTighe, 2007). However, if teachers could have presumed that all involved in the

discussion of grading practice reform came from a place of love for kids and caring about

their future, they might have begun to see that "suggestions of reform are not a criticism

of the past but a hope for the future" (Reeves, 2011, p. 79). Allen (2005) wrote, "Because

grading is something that has been done to each of us during our many years as students,

it is hard to change the invalid grading schema that has become embedded in our minds"

(p. 218).

Hargreaves and Fullan (2013) likened teachers' ability to exercise personal

judgment in grades to the legal concept of decisional capital. Decisional capital referred

to professionals, such as judges or teachers, wherein the capacity to judge was developed

over time; it was based on situations and circumstances where the evidence and the

answers were not "incontrovertibly clear" (Hargreaves & Fullan, 2013, p. 37). The

literature showed that this notion of decisional capital was a common sentiment

expressed by teachers and might have served as a reason as to why some teachers chose

not to follow expert recommended grading practices (Allen, 2005; Guskey, 2004;

Yesbeck, 2011). Rather, decisional capital inclined teachers to formulate their grading

practices around their classroom experiences (Allen, 2005; Guskey, 2004; Yesbeck, 2011).

Many experts stated that teacher education programs also contributed to the issue of varying grading practices, because they lacked clarification of grading practices and policies in their courses (Allen, 2005; Brookhart, 1994; Stiggins, 1999). Guskey (2004) cited that novice teachers had little knowledge and received limited training in effective grading methods and practices. Zoeckler (2007) added, "It is clear that teachers' understandings of what grades mean and how they are meant to be understood is an area in need of exploration" (p. 86). According to Allen (2005), less than one-third of teacher training programs required pre-service teachers to complete an assessment course. Furthermore, teacher education programs emphasized methods of instruction, as opposed to the development of appropriate assessment measures and the contributing factors to consider when grading (Yesbeck, 2011).

Allen (2005) maintained that, for years, experts communicated all of these problems to K-12 educators, teacher preparation programs, and teacher development trainers and planners with little to no impact on practice. This issue was further evidenced in research conducted by Lomax (1996), wherein he tracked a group of pre-service elementary teachers throughout the tenure of their training program and noted grading proved to be the teachers' greatest challenge. Allen (2005) purported, "One of the goals of a teacher education program should be to prepare pre-service and in-service teachers to develop effective methods to assess students and to communicate clearly and accurately through their grading practices that assessment to others" (p. 220).

However, experts and teachers disagreed about the overall quality of teacher

preparation programs. Despite experts' reservations about the lack of instruction about best-grading practices received by education students, most teachers reported they were "generally satisfied with their preparation programs" (Feistritzer, 2011, p. 24). Of the teachers surveyed, 24% of the teachers surveyed indicated that they had rated their preparation program as "Excellent," 41% said "Very Good," and 24% said it was "Good." Only 1% rated their program as "Poor" (Feistritzer, 2011, p. 24).

Research conducted by Feistritzer (2011) found that one-third of first-time public school teachers hired since 2005 obtained their positions after completing an alternate program other than a college campus-based teacher education program (p. ix). Like teachers who earned their certifications via the traditional route, alternative program teaching graduates received limited training in regards to best-grading practices. This was significant in that these alternative programs were producing approximately 60,000 new teachers per year (Feistritzer, 2011, p. ix).

Therefore, because most teachers received only a small amount of formal training in effective grading methods, and others received none at all, teachers usually drew from their personal experiences as students, that they regarded as being fair and reasonable in determining the grading practices they employed (Bailey, 2012; Frary et al., 1993; Guskey, 2006, 2009; Guskey & Bailey, 2001; Leyrer, 2015). In other words, as Guskey (2004) pointed out, "Most teachers do what was done to them" (p. 49). Because recollections of these experiences varied among teachers, the grading practices they employed varied, as well (Guskey, 2006).

Despite their critical assessment role and widespread use, grades and grading practices varied significantly by teacher, grade level, subject, and school (Leyrer, 2015).

According to McMillan (2008), teachers used their intuition, subjective evaluations, personal values, beliefs, and philosophies in grading. As a result, grades were often regarded by stakeholders as arbitrary (Dockery, 1995) and regarded by measurement experts as invalid (Wormeli, 2006). In the words of Marzano (2000), "Grades are so imprecise they are almost meaningless" (p. 1).

The act of grading had long been a significant responsibility for educational professionals (Guskey, 2004). Chiekem (2015) asserted, "Assigning grades is probably the most important measurement decision that classroom teachers make" (p. 24). Brookhart (2011) wrote that when she talked to teachers about their grading practices, "feelings often run high" (p. 10). "Teachers tend to assume that others agree with their positions, but in fact I hear a range of opinions" (Brookhart, 2011, p. 10).

Teachers' employment of varying grading practices was not only problematic to the integrity of educational professionals, but to students, as well. Teachers' grading practices resulted in the assignment of grades that affected students' opportunities and, ultimately, their futures (Leyrer, 2015). Several institutions used a student's grades in making important educational, financial, and career decisions (Nitko, 2001). As stated by Johnson and Johnson (2002), "The power of grades to impact students' future life creates a responsibility for giving grades in a fair and impartial way" (p. 249).

Throughout the course of their education, students learned to interpret the grades they received from teachers as a method of communication that represented how much they knew and indicated the quality of their academic work (Leyrer, 2015). Students associated specific grades as indicators of learning progress and academic standing that became increasingly important, due to their ability to determine academic placement,

GPA, and college admissions. As and Bs were believed to have been good grades, Cs were so-so grades, and Ds and Fs were poor grades (Leyrer, 2015). Grades, particularly at the secondary level, served as currency in the educational marketplace (Brookhart, 1993). Expanding on Brookhart, Yesbeck (2011) stated, "Grades have become the standard by which many students, schools, and programs are compared" (p. 24).

Webster (2011) synthesized research and concluded, "Traditional grading practices foster a performance-focused orientation by rewarding success compared to peers and by highlighting ability differences and levels of students" (p. 44). This categorization of students caused some educational professionals to question the value of assigning grades. As stated by Brookhart (2011), "More and more educators are beginning to question traditional grading practices that were developed to sort students into learners and non-learners, not to support learning for all" (p. 10). Furthermore, some educators argued that the use of extrinsic rewards to reinforce the intrinsic process of learning taught students to care more about their grades than knowledge (Edwards, 1999). This sentiment dated back to research conducted in 1911 that raised a similar concern that grading reduced students' intrinsic interest in learning (Colvin, 1912).

Critics of the traditional grading system maintained that grades, and the competition they created among students, could have resulted in psychological harm to students (Kohn, 1999). Sinclair and Ghory (1987) found that, due to the belief that grades were interpreted as measures of a student's merit, many youths felt marginal to the central school population, in part because they were receiving messages in the form of failing grades that they did not belong in school. Additionally, Allen (2005) maintained that, when grades under-valued students' knowledge and skill level due to the nature of

the factors included by a teacher when grading, it could have detrimental effects on the

student's self-esteem and future opportunities.

As a result, some critics suggested that the act of assigning grades to students be

abolished altogether (Kohn, 1994). However, Schneider and Hutt (2013) stated that

opponents of grades and their criticisms "have never occupied more than the margins of

the discussion over grading and they have done little to dislodge formal marking

systems" (p. 2). Wiggins (1996) echoed this sentiment, arguing that stakeholders would

be best served by trying to make grades better, as opposed to merely abolishing them.

According to Feldmesser (1971), grades could have positively affected students'

motivation to learn and self-perception. However, students who received low grades

often experienced apathy and generally withdrew from class work (Guskey, 1994). This

was especially problematic, because children who lacked confidence in their ability could

have developed "learned helplessness" that would have hindered them from continuing to

try after experiencing failure (Webster, 2011, p. 2). Reeves (2008) declared, "If you

wanted to make just one change that would immediately reduce student failure rates, then

the most effective place to start would be challenging prevailing grading practices" (p.

85).

When grading practices varied considerably from one teacher to another, grades

could have resulted in tension between teachers and students (Randall & Engelhard,

2009). A survey conducted in 1897 found that high school students had "long pent-up

indignation against the grading system" (Hyde, 1897, p. 92). Research conducted by

Holmes and Smith (2003) surveying students' opinions about how teachers graded their

assignments revealed that students were aware of both the importance of grades and that

grading practices varied significantly from teacher to teacher. Students reported that the issue of grade fairness was a concern, because they felt grades were inconsistent with other students' similar work that earned different grades. Students also reported teachers did not provide sufficient feedback regarding an explanation of the grade they received and/or how to improve the grade. Therefore, as stated by Reeves (2008), "the difference between failure and the honor roll often depends on the grading policies of the teacher" (p. 85).

According to Jongsma (1991), grades could "misinform and deceive" (p. 318), because the act of assigning a student one letter grade implied a level of evaluation precision that was nearly impossible. This, coupled with teachers' varying grading practices, was likely to frustrate stakeholders and serve as a source of conflict among relevant parties (Randall & Engelhard, 2009). Within the traditional 'A-F' letter grading system used by most schools, students grew to understand they are assigned various numbers of points by teachers for differing levels of achievement ("Effective Grading Practices," 2011). They also grew to understand they must accumulate a specific number of points to achieve high grades ("Effective Grading Practices," 2011). Because letter grades were associated with specific numerical values, Erickson (2011) argued, "At a very young age, most students learn the point system and how school can be about the accumulation of points, not the accumulation of knowledge and skills" (p. 43).

When discussing the subject of grades, the terms 'measurement, assessment and grades' were often used interchangeably, because they all described evaluation, to varying degrees (Yesbeck, 2011). Airasian (1997) stated that assessment was a way to help stakeholders in their decision-making process by collecting, synthesizing, and

interpreting information. McMillan (2008) stated that measurement was a process employed to determine the degree to which something was demonstrated and was given a value relative to a scale. Grades should have been the result of clear measurement of the best a student was capable of (Yesbeck, 2011). However, because grading tended to be based on professional judgment, the practice varied significantly (McMillan, 2008). A major consequence of teachers including non-academic factors in grades was it resulted in an ineffective communication of a student's academic abilities (Tomlinson & McTighe, 2006).

Researchers found that ineffective grading practices, such as including non-academic factors in grades, also contributed to the high school failure rate (Reeves, 2008). As stated by Reeves (2008), grading practices had a direct effect on the grades students received from teachers; so, it was of the utmost importance that schools carefully considered which grading practices best assessed and reported student performance. Experts suggested that better grading practices might have helped to reduce the more than 330 billion dollar annual cost of high school failure ("Effective Grading Practices," 2011).

In addition, researchers found that ineffective grading practices were problematic, because they contributed to the discrepancy between grades and student performance scores on state assessments (Guskey, 2006). Simply stated, students receiving high grades should have been scoring higher on state assessments than they were. Teachers and policy makers disagreed as to why students receiving high grades were not consistently top performers on state assessments. Teachers blamed test makers, citing poorly constructed assessments, while policy makers blamed teachers, arguing their grading

practices were biased and subjective, resulting in inaccurate measurement of student

achievement (Guskey, 2006).

O'Connor (2009) expanded upon Guskey's findings by contending that grades

should have been considered ineffective when they failed to satisfy the following four

conditions: (a) They were accurate; (b) they were consistent; (c) they were meaningful;

and (d) they were supportive of learning. Brookhart (2009) believed that, ideally, grades

would have been replaced with discussions of student performance, wherein individual

strengths and weaknesses were emphasized in order to guide students through the process

of improvement, until a task or skill was mastered. However, Wiggins (1996) cautioned

educators that any attempts to abolish letter grades would likely lead to political battles,

because parents would have been suspicious of individuals who sought to challenge a

120-year-old system they thought they understood, despite any reservations they might be

harboring about the validity of grades. In other words, comfort and familiarity would

have trumped concerns about fairness and validity. Wiggins (1996) further explained that

opponents of traditional letter grades often failed to comprehend that the symbols

themselves were actually not the problem; instead, the absence of consistent and clear

points of reference in using the symbols was the problem.

Some districts adopted grading policies in effort to address inconsistent grading

practices and limit subjectivity (Polloway et al., 1994; Strein & Meshbesher, 2006).

Grading policies were the umbrella under which grading practices were created

(Yesbeck, 2011). However, Austin and McCann (1992) found that school board policies

varied considerably and may or may not have included information about the following

components: (a) the purposes of grades; (b) audience for grades; (c) the criteria for

calculating grades; (d) best-grading practices; (e) authority of school board policies; (f) quantification of building-level support regarding grading practices; and (g) professional development. Furthermore, Cizek et al. (1996) found that "many teachers do not share a common knowledge or understanding of district policies on grading where such policies exist" (p. 172).

While effective grading practices were proposed, they were not adopted on a widespread scale (Cross & Frary, 1999; O'Connor, 2007). Teachers continued to employ a system of varying factors, based on personal experiences, despite established grading practice recommendations (Brookhart, 1994; McMillan, 2001). Furthermore, researchers found that even in districts possessing policies to guide teachers' grading practices, many teachers ignored the policies and included factors outside of the criteria established within them (Bruckman, 2010; Reeves, 2008; Strein & Meshbesher, 2006) - especially when calculating grades for borderline students (Randall & Engelhard, 2010). As stated by Reeves (2008), "Neither the weight of scholarship nor common sense seems to have influenced grading policies in many schools" (p. 85).

The integration of technology in schools and the emphasis placed upon it throughout the past two decades resulted in the replacement of the printed report card with an electronic report card (Tetreault, 2005). Many schools provided their teachers with computers to access electronic grading systems that enabled them to use paperless reporting (Tetreault, 2005). The many advantages of using electronic grading systems were that they were created to compile grades accurately, compute averages, weigh scores according to teacher instruction, flag students with specific characteristics, print reports for individuals and/or groups, and allow stakeholders electronic access to grades

(Yesbeck, 2011). However, because teachers' grading practices differed, electronic grading systems presented challenges, such as how to record missing work, how zeroes affected the average, and how different assignments were weighted, according to each teacher's preference (Yesbeck, 2011). Teachers who included non-academic factors in grading typically considered electronic grading systems too rigid, because such systems relied upon objective mathematical calculations instead of subjective concepts, such as fairness and truth (Guskey, 2002). Although, some electronic grading systems were equipped with an option allowing teachers to create a category for non-academic factors (Yesbeck, 2011).

Guskey (2002) warned teachers that electronic grading systems were not the end all solution to the problem of teachers mixing non-academic factors with academic factors in grading. According to Guskey, computerized grading systems were synonymous with a myth of objectivity based on teachers' assumptions that, as long as the mathematical computations were correct and all students were treated the same, the grades assigned in such systems were accurate and judicial. This perspective failed to take into account that numerical precision differed from evaluative fairness, honesty, or truth (Guskey, 2002). While such systems may have been useful for the purpose of simplifying record keeping, they did not guarantee that the teacher was unable to manipulate the settings in a manner that undermined accurate assessment of a student's academic performance devoid of bias.

An era of increased student accountability and high-stakes testing was emerging wherein the popularity of standards-based grading and the use of course objectives to determine student mastery was increasing (Tomlinson & McTighe, 2006). However, the

implementation of various grading policies and expert recommended grading practices in

the 21st century resulted in an overall feeling among teachers of reduced autonomy and

empowerment (Leyrer, 2015; Mulford, 2003; Stacy, 2013). Individual, personalized

grading practices were, in the opinion of several teachers, "the last bastion of teacher-

controlled policy" (Leyrer, 2015, p. 2). While educational policy changes at the state and

federal levels focused on the attainment of mandates of student achievement, more

research focused on teachers' grading practices (Bailey, 2012).

It was the responsibility of all teachers to exercise grading practices that resulted

from careful reasoning about which specific factors should be included in determining

students' grades (Guskey, 2002). It was imperative that teachers understood how

academic and non-academic factors should be summarized and what format should be

used to report them. Regardless of the computer program or various other tools available

to assist teachers with calculation grades, the most essential component of the process

was the teacher, because he/she was the ultimate adjudicator of accuracy and fairness

(Guskey, 2002). Therefore, it was essential that teachers' grading practices reflected such

ideals.

Research showed that grading was considerably influenced by the values and

beliefs of the teacher (Buzzelli & Johnston, 2002). As stated by Hawkins (2010), no other

factor influenced teachers' grading practices more than the individual teacher's personal

values and beliefs. Buzzelli and Johnston (2002) maintained that students who better

understood their teacher's values and beliefs were more likely to have purposely

demonstrated characteristics consistent with such. Doing so made these students more

likely to receive higher grades than their peers who did not understand their teacher's

values and beliefs and/or failed to demonstrate characteristics consistent with such. Brookhart (1993) also emphasized the significance of the individual teacher's values and beliefs in grades. She found that non-academic factors were often used by teachers who regarded the consequences of grades as being more important than grades' ability to have communicated information about student learning.

However, despite personal and professional histories and long traditions of varying grading practices used by teachers, educators could have changed their grading practice guidelines and policies to more effective ones - under certain influences (Dyd, 2012). While teachers' grading practices could have been influenced by their personal philosophy of teaching and learning (Tomlinson, 2001), they could have also been influenced by official grading policies and perceived and actual consequences of grades (Randall & Engelhard, 2009).

While the initial act of changing grading systems could have been overwhelming for school leaders and administrators, "the benefits are so great that it's worth doing" (Reeves, 2008, p. 87). Reeves (2011) further stated that such benefits were often the result of low rates of student failures leading to reduced discipline problems, increased college credits, more elective courses, improved teacher morale, fewer hours of Board of Education time diverted to suspensions and expulsions, and added revenues for the entire system, based on a higher number of students continually enrolled in school (p. 79).

Therefore, the benefits resulting from grading practice reform could have been significant for schools and districts as a whole (Heflebower, Hoegh, & Warrick, 2014). Although grading practices constituted only one of several pieces of the educational process (Reeves, 2011), Fullan (2010) maintained that even when just one piece was out

of alignment, every other part of the system felt it. And, as stated by Reeves (2011), this was felt "for good or ill" (p. 79).

Academic Factors

The one standard factor accounted for in all grading systems was that of student academic achievement (Wormeli, 2007). Experts agreed the main purpose of grades was to provide and report achievement information (Bailey & McTighe, 1996). Non-educators also agreed (Simon & Bellanca, 1976). As stated by Fleming (1999), "Quality teaching has few measurable outcomes and so too does quality learning. The universal measure for the latter is the test, the assignment, the project, the event, the artifact or the examination" (p. 84). Therefore, it was essential that the specific factors teachers included in their grading practices were based off of valid measures of academic achievement alone (Chiekem, 2015). These factors were referred to as academic factors.

Academic factors consisted of those considered to have related to a student's learning achievement, as opposed to factors that related to their behavior (Guskey, 2002). Experts recommended that grades should be based only off of academic factors (Wormeli, 2006) with non-academic factors reported separately (Allen, 2005; Cross & Frary, 1999; Guskey, 2004). However, while teachers maintained that they believed achievement was the most important factor in a grade (Allen, 2005; McMillan, 2007; O'Connor, 2009; Tomlinson, 2005; Webster, 2011), their grades were rarely the result of solely academic factors (Brookhart, 1991). Instead, Wormeli (2006) stated that teachers mistakenly used non-academic factors in their grading practices to promote ideals of student accountability.

Stiggins (2005) defined assessment as "the process of gathering evidence of student learning to inform instructional decisions" (p. 5). He cautioned teachers, "This process can be done well or poorly" (p. 5). But to be considered efficient in the classroom, teachers "must be able to do it well" (p. 5).

There were three types of assessment - diagnostic, formative, and summative. Diagnostic assessment, also known as pre-assessment, was designed to help teachers evaluate students' prerequisite knowledge prior to beginning a unit or lesson (Tomlinson, Moon, & Imbeau, 2013). They assisted teachers in knowing "when to reteach, when to move ahead, and when to explain or demonstrate something in another way" (Tomlinson, 2007, p. 11). Simply stated, diagnostic assessments served as tools for "the beginning of better instruction" (Tomlinson, 2007, p. 11).

Formative assessments pertained to the collection of evidence that helped students make informed decisions about how to make improvements regarding their learning (Brookhart, 2009; McMillan, 2008; O'Connor, 2012; Popham, 2008; Yesbeck, 2011). They could have been used to provide direction for the improvement of individual students or for entire classes (O'Connor, 2012). Formative assessments ranged from instructional questioning to classwork assignments and provided teachers with information about students' current level of understanding (Yesbeck, 2011). They were designed to function as feedback that assisted teachers in making instructional decisions (Brookhart, 2009).

Summative assessment referred to evaluations that measured student work over a period of time in an attempt to collect evidence in reporting individual student grades (Brookhart, 2009). Summative assessments consisted of unit and standardized tests, as

well as any assignment factored into the final course grade (Brookhart, 2009). They were designed to measure students' comprehension of material over a period of time and the data were especially useful to schools when making decisions about instructional programs (Yesbeck, 2011).

While all types of assessment were useful in gathering information about student achievement, only achievement information obtained from summative assessments should have been considered when calculating academic grades (Brookhart, 2009; McMillan, 2008; O'Connor, 2012). However, research showed that some teachers did not clearly understand the difference between formative and summative assessments (Yesbeck, 2011). As stated by Cizek et al. (1996), "It seems that classroom assessment practices may be a weak link in the drive toward improving American education" (p. 162). Throughout a given grading period, most teachers averaged a student's performance on tests and assignments (O'Connor, 2009). In doing so, they were inaccurately combining formative and summative assessments (Airasian, 2005). When "three components - measurement, assessment and grading - are aligned, then the grade reflects a true indication of the student's understanding on an assessment that was developed based on measurement recommendations" (Yesbeck, 2011, p. 23).

Regardless of the type of assessment, Stiggins (2005) argued that students should play an active role in the grading process. O'Connor (2007) stated that to promote achievement, teachers should have allowed students to play key roles in both the assessment and the grading process. Stiggins (2005) suggested three ways teachers could have accomplished this: (a) Involve students in the construction of assessments and in the development of criteria for success; (b) have students keep personal records and/or

portfolios of their achievement and growth throughout a course; and (c) allow students

opportunities to communicate their achievement such as being involved in parent

conferences.

Non-Academic Factors

Non-academic factors were those considered in grading practices related to

student behaviors, work habits, and attitudes (Brookhart, 2009). A synthesis of research

indicated they included the following: organization, ability, aesthetic appearance of work,

attendance, behavior, difficulty level of an assignment, effort, attitude, motivation, extra

credit, homework, completion, improvement, participation, punctuality, and

responsibility (Bailey, 2012; Brookhart, 1994, 2009; Guskey, 2004; McMillan, 2001,

2002; O'Connor, 2007). The use of zeroes as a punishment for late or missing work was

regarded by researchers as a final non-academic factor (Canady & Hotchkiss, 1989;

Guskey, 2004; Stiggins & Duke, 1991).

Measurement experts strongly cautioned against the use of non-academic factors,

because they compromised the validity of grades when combined with academic factors

(Bailey, 2012; Brookhart, 1993, 1994; McMillan, 2001, 2002; McMillan & Lawson,

2001; McMillan et al., 2002; McMunn, Schenck, & McColskey, 2003; Stiggins et al.,

1989; Wormeli, 2007). McMillan (2001) even went so far as to refer to non-academic

factors collectively as "academic enablers," because they were not indicative of academic

achievement and only served to inflate a student's grade (p. 25). Guskey (2011) likened

teachers' current practice of combining academic and non-academic factors into one

grade to doctors combining unrelated health measures into a single score. Guskey wrote,

If someone proposed combining measures of height, weight, diet, and exercise

into a single number or mark to represent a person's physical condition, we would

consider it laughable. How could the combination of such diverse measures yield

anything meaningful? Yet every day, teachers combine aspects of students'

achievement, attitude, responsibility, effort, and behavior into a single grade that's

recorded on a report card- and no one questions it. (p. 19)

Teachers often cited that they included non-academic factors in grades because it

was part of teaching students to be accountable or responsible, so they were prepared for

life in the real world outside of school (Wormeli, 2007). This practice went against expert

recommendation, which strongly rejected the idea of teachers using grades in a punitive

nature (Cizek et al., 1996; Wormeli, 2007). Cizek (2003) warned,

> When a teacher uses grades as punishment for student behaviors, the teacher
>
> establishes an adversarial relationship in which grades are no longer meaningful
>
> to students as indicators of their accomplishments. The punitive use of grades
>
> only increases the likelihood that students will lose respect for the evaluation
>
> system." (p. 43)

Wormeli (2006) further cautioned,

> A grade is supposed to provide an accurate, undiluted indicator of a student's
>
> mastery of learning standards. That's it. It is not meant to be a reward, motivation,
>
> or behavioral contract system. If the grade is distorted by weaving in a student's
>
> personal behavior, character, and work habits, it cannot be used to successfully
>
> provide feedback, document progress, or inform our instructional decisions
>
> regarding that student- the three primary reasons we grade. (p. 19)

Regrettably, however, many grades did not match this description (Heflebower et al.,

2014, p. 4). Canady and Hotchkiss (1989) revealed that more than a third of teachers believed grades served as a means of punishment. This occurred despite the evidence that showed grades were not effective when used as such.

Wormeli (2007) purported that teachers placed "too much under the banner of a grade that has little or nothing to do with the achievement that a grade is supposed to represent" (p. 75). Instead, Wormeli stated, "A grade should serve as an accurate, undiluted indicator of a student's mastery of learning standards" (p.75). Experts argued that grades resulting from the combination of academic and non-academic factors were inaccurate measures of student mastery because they were contaminated with non-academic information that would have only served to dilute their meanings (Allen, 2005; Greville, 2009; Shanahan, 2011).

Guskey (2006) found that high school teachers considered several different academic and non-academic factors when grading. These factors included, but were not limited to, "major exams or compositions, homework completion, class quizzes, class participation, reports or projects, work habits and neatness, student portfolios, exhibits of student work, effort, attendance, laboratory projects, punctuality of assignment submissions, student notebooks or journals, class behavior or attitude, classroom observations, oral presentations, and progress made" (Guskey, 2006, p. 671). Guskey (2006) found that most teachers based grades on as little as two of these factors or as many as 16. He further stated this variation existed even among teachers who taught in the same school.

When teachers mixed academic and non-academic factors into a single letter grade, it quashed the grade's ability to clearly communicate any one aspect of a student's

education (Guskey, 2001, 2009, 2013; Leyrer, 2015; Marzano, 2000, 2006, 2010; Stiggins, 2001). This practice resulted in grades that were inaccurate measures of student mastery after they had been distorted with non-academic factors, such as attendance, effort, behavior, and ability (Allen, 2005; Greville, 2009; Shanahan, 2011). For this reason, non-academic factors were also referred to as academic enablers (Bailey, 2012; McMillan, 2001, 2002). Their inclusion in grading allowed greater potential for discrepancies between teachers' grades due to subjective teacher bias (Bailey, 2012; Cross & Frary, 1999).

Winger (2009) stated that when non-academic factors were combined with academic factors in grades, it sent the message to students that compliance was the priority - not learning. As a result, students often began manipulating the learning process - a phenomenon Scriffiny (2008) referred to as "playing school" (p. 71). An example of this was when students who learned very little were able to rely on non-academic factors, such as homework and extra credit, by which they improved their grade. However, such students often continued to struggle with academic factors throughout their remaining years of school (Scriffiny, 2008).

Guskey (2006) wrote that even when teachers disclosed the weighting proportions they used when combining specific academic and non-academic factors in grading, and used computerized grading programs to ensure accurate calculations, the meaning of the letter grade was still compromised. Guskey argued that, even when taking such precautions, the final grade remained "a confusing amalgamation that is impossible to interpret and rarely presents a true picture of a student's proficiency" (p. 671-672); thus, having made the grade unfair and inaccurate.

Randall and Engelhard (2009) found that teachers generally assigned day-to-day grades based on academic factors; but, then included non-academic factors when determining final grades. As a result, Conley (2000) found that there was little relationship between the final grade the teacher assigned and the proficiency of the student. It was, therefore, ironic that teachers' varying grading practices, which resulted in inconsistencies in grades were tolerated in the field of education, when a similar occurrence would probably never have been tolerated among professionals in other fields, such as sports or medicine (Reeves, 2008).

Experts argued that the practice of including non-academic factors in grades allowed greater potential for discrepancies between teachers' grades, due to subjective teacher bias (Bailey, 2012; Cross & Frary, 1999). Therefore, experts suggested that teachers report non-academic factors separately from academic factors, such as in a standards-based grading system. This helped to ensure that the course grade only consisted of academic factors, but reported supplemental information pertaining to non-academic factors (Guskey, 2002). Costello and McKillop (2000) reported an increasing number of schools began to use grade reporting systems that indicated a student's work habits in addition to grades. Schools using these systems were actually able to provide stakeholders with both more information and more accurate information. This was because such schools were able to clearly pinpoint a student's strengths, as well as the areas in which they needed to improve, while simultaneously preserving the validity of the grade (Guskey, 2002).

However, the debate among experts and teachers as to whether non-academic factors should have been included with non-academic factors in grading remained one of

the most contentious aspects regarding the topic of grading practices (Leyrer, 2015). Researchers found many teachers had a history of including several non-academic factors in their grades (Bailey, 2012; Brookhart, 1993, 1994; McMillan, 2001, 2002; McMillan & Lawson, 2001; McMillan et al., 2002; McMunn et al., 2003; Stiggins et al., 1989; Wormeli, 2007). The reasons cited by teachers for including non-academic factors in their grading practices often varied by grade level. For example, previous studies indicated that elementary teachers were less likely to assign their students failing grades than were middle school teachers, because elementary teachers placed greater emphasis on effort (Bursuck et al., 1996; Randall & Engelhard, 2009). Middle school teachers had more stringent grading practices, because they felt internal and external pressures to prepare their students for the increased rigor of high school (Randall & Engelhard, 2009).

Randall and Engelhard (2009) suggested additional possible reasons as to why this emphasis on effort occurred more at the elementary level than the middle school level. They stated that because elementary teachers spent more time with their students, they could "feel compelled by their very nature to nurture and protect the self-esteem of their students" (p. 184). Or, perhaps elementary teachers felt pressure to work harder, so their students did not fail and any such failure by their students would have constituted a negative reflection of their abilities as a teacher. This variation in grading practices from one level to the other may have confused students about the meaning of grades (Randall & Engelhard, 2009).

One reason teachers cited for including non-academic factors in grades related to student motivation and behavior. Ideally, teachers preferred that all students possess a sophisticated level of intrinsic motivation regarding school (Guskey, 2004). However, in

reality, this was certainly not the case. Researchers found intrinsic motivation to learn

declined throughout the middle school years (Anderman, Maehr, & Midgley, 1999;

Gottfried, 1985). As a result, most teachers reported that grades served important roles in

regard to the amount of effort students put forth (Chastain, 1990; Ebel, 1979; Guskey,

2004).

Although teachers might not have been outright in admitting that the grades they

gave students served as punishment, researchers found that teachers mistakenly used

grades as incentives for students to learn (Guskey & Bailey, 2001). Guskey (2002) stated

that a typical, yet detrimental, grading practice pertaining to non-academic factors that

was commonly used by teachers was the act of lowering students' grades because of poor

behavior. Researchers found that teachers often lowered grades due to non-academic

factors, such as disruptive behavior, because they felt doing so helped to motivate

students to complete their work and not misbehave (Cross & Frary, 1999; Johnson,

McGue, & Iacono, 2005). However, Wormeli (2006) clarified that grades were not

supposed to serve as a reward, motivation, or behavioral enforcement system, despite the

fact that many teachers used them to punish bad behavior or to reward/motivate for good

behavior.

Classroom disruptions, tardiness, attendance, academic dishonesty, and attitude all

fell under the umbrella of behavioral infractions that were heavily weighed by teachers

when calculating grades (Guskey, 2002). Guskey (2002) maintained that such a reduction

violated the purpose of providing a summary judgment of student achievement and

blatantly miscommunicated academic assessment information. He stated that while

students' behavioral infractions should not have gone unnoted by teachers, "it's clear that

they are not part of the evidence that shows what these students have learned and are able

to do" (p. 780). Guskey (1996) likened teachers' abilities to reduce grades for

misbehaving students as a "weapon of last resort" (p. 18). Guskey wrote,

> Students who do not comply with their requests suffer the consequence of the
>
> greatest punishment that teachers can bestow: a failing grade. Such practices have
>
> no educational value, and, in the long run, adversely affect students, teachers, and
>
> the relationship they share. (p. 18)

Another reason why teachers included non-academic factors in grades related to

communication. Many teachers used non-academic factors in their grades to

communicate big-picture messages to students (Allen, 2005) about level of expectation,

encouragement, and disappointment, in addition to level of academic achievement

(Zoeckler, 2007). "These messages pertain to individual teachers' ideas of what is right

and wrong and what is good and bad and are significantly figured into their evaluation of

student work and the grades that they assign to their students" (Zoeckler, 2007, pp. 97-

98). However, McMillan (2012) stated that this practice contradicted the belief that

grades should have been used for student reaction and reflection regarding self-regulation

of learning. "If students are to use their grades in an evidentiary process to regulate even

more learning, the grades need to be evidence of learning" (McMillan, 2012, p. 45).

Teachers also felt political pressure that impacted their grading practices

(Webster, 2011). As a result of legislation, such as the 2002 No Child Left Behind Act,

teachers felt the pressure of increased testing demands and accountability standards

(Linn, Baker, & Betebenner, 2002). Dunn and Allen (2009) attributed the heightened

testing requirements to a shift in accountability that caused schools, and teachers in

particular, to feel an even greater sense of accountability for the growth of their students. In addition to legislative, institutional, and societal notions of increased accountability for student achievement, there were also increasing expectations for teachers to produce data that substantiated the efficiency of their instruction (Slavin, 2007).

Similarly, Brookhart (1993) found that teachers often included conceptions of student effort when grading, because they were concerned with the social consequences of their grading practices. However, Guskey (2002) argued, if teachers changed the way they reported non-academic factors by reporting them separate from academic factors, they could have still satisfied areas, such as motivation, self-esteem, and social consequence, while maintaining the intended purpose of grades. As stated by Scriffiny (2008), while teachers were not able to control many factors in a school setting, such as funding, teaching assignments, or other important issues, teachers could control how they assessed student learning.

Reeves (2008) purported that one of the most commonly used, yet ineffective, grading practices was the assignment of zeroes for missing work. Research showed that while few teachers believed that grades should have been used punitively to punish students for a lack of effort or responsibility, many teachers assigned zeroes for work submitted late, missing, or incomplete (Stiggins & Duke, 1991; Guskey, 2002). While most teachers recognized that giving a student a zero for poor behavior punished them academically, most teachers still believed that such a punishment was both justified and deserved (Guskey, 2004). In this way, teachers used zeroes as their "ultimate grading weapon" (Guskey, 2004, p. 50). However, this practice went against expert

recommendation that grades should have reflected how well students learned (Guskey, 2002).

When used in a punitive manner, zeroes represented an instrument of control that teachers used over students (Guskey, 2004). Teachers used zeroes in grading to punish students' behavioral offenses, such as being irresponsible, not putting forth sufficient effort in school work, disobeying the teacher's warnings, violating the teacher's expectations, and missing assignment deadlines (Canady & Hotchkiss, 1989; Guskey, 2004; Stiggins & Duke, 1991). As stated by Guskey (2004), "The threat of a zero- and the resulting low grade- allows teachers to impose their will on students who otherwise might be indifferent to a teacher's demands" (p. 50). Expanding on Guskey (2004), Reeves (2008) stated, "Defenders of the zero claim that students need to have consequences for flouting the teacher's authority and failing to turn work in on time. They're right, but the appropriate consequence is not a zero; it's completing the work" (p. 85).

While some teachers defended their practice of giving zeroes as having been based on the concept that they could not give credit to students for work that was either lacking sufficient effort or not turned in at all, there were recommended alternatives (Guskey, 2002). Guskey proposed that in lieu of assigning zeroes teachers assigned an 'incomplete' or required a student to attend a special study session after school or over the weekend in order to regain their credit. Reeves (2008) concurred, suggesting that work could have been completed before or during school, during study halls, at lunc, or other settings.

Another alternative to giving a zero was assigning an 'I' for insufficient evidence (O'Connor, 2007). These alternatives taught students responsibility and consequence, because they were held accountable for doing satisfactory work (Guskey, 2002). Teachers would have found that students would have risen to the high expectations surrounding such policies, because they knew that teachers were serious about enforcing student responsibility (Guskey, 2004). At the same time, teachers would have preserved the integrity of their grades, because they would have remained an accurate reflection of what students had actually learned (Guskey, 2004).

Some schools may have objected to implementing these suggested alternatives due to increased financial costs they thought they would have incurred. After all, these alternatives did require schools to acquire or allocate necessary funding to secure additional pay for resources, such as teachers and classroom space (Guskey, 2004). However, schools that implemented such policies found they actually saved money in the end. Due to the immediate nature of the consequences, students would quickly remedy learning or behavioral issues before they developed into major problems (Guskey, 2004). This served as an investment in resources, because heading off the issues in the beginning saved money down the line.

A single zero could have a particularly detrimental effect on a student's grade when combined with the commonly used practice of grade averaging (Guskey, 2013). In a grade averaging system, when a teacher entered a zero for a student, it considerably skewed their course grade by distorting the average. As stated by Zoeckler (2007), the major concern about grade averaging was that doing so interfered with the truthfulness of grades as an indication of achievement. Many experts regarded grade averaging as an

unfair practice, because in order to recover their grade from a zero, a student would have

to perform perfectly on several subsequent assignments (Guskey, 2013). Therefore, the

assignment of a single zero had the potential to condemn a student to failure (Guskey,

2004) with virtually no chance of recovery (Guskey, 2013).

Guskey (2004) suggested a relatively simple action schools could have taken to

mitigate the negative impact of zeroes on students' grades was to alter the grading scale

they were using. This entailed shifting from percentage grading to whole number grading.

For example, this would have looked like going from earning A (representing 90 to

100%), B (representing 80 to 89%), C (representing 70 to 79%), D (representing 60 to

69%), and F (representing 0 to 50%) to A (representing 4), B (representing 3), C

(representing 2), D (representing 1), and F (representing 0). This shift would have

continued to allow teachers to assign zeroes for unsatisfactory or missing work; but, it

would have lessened the effect of receiving a zero. According to Guskey (2004), this

policy did not address the problem of ensuring students' grades accurately represented

their learning, but it did help protect students from the considerably detrimental impact of

earning a zero in a percentage grading system.

Reeves (2008) bluntly stated the use of zeroes for missing work "should be

labelled as toxic" (p. 86) due to the crippling effects it could have had on a student's

grade. Reeves (2010) likened the practice of assigning zeroes within a 100-point grading

scale as a punitive grading measure to administering an academic death penalty. Reeves

(2010) wrote,

Assigning a zero on a 100-point scale is a math error; it implies a 60-point

difference between the 'D' and 'F.' . . . It makes missing a single assignment the

academic death penalty. It's not just unfair - it is not mathematically accurate.

(p. 78)

Like late grade deductions, zeroes also represented teachers' attempts to affect students'

behaviors by statistically incorporating punitive measures into their grading practices

(Dueck, 2014).

As a result, states, such as Texas and Nevada, had gone so far as to pass

legislation stating that the lowest percentage teachers could assign to students was 50

rather than zero (Montgomery, 2009; Richmond, 2008). School districts adhering to

minimum-grade legislation/policies operated under the premise that they would not give

credit to students when no credit was due because a percentage grade of 50 or less was

still considered a failing grade in almost every school (Guskey, 2013). Other districts

used minimum-grade policies in attempt to mitigate the confounding effects of a zero

used in a percentage system (Guskey, 2013).

Best Practices for Grading

Reeves (2008) emphasized to educational professionals the importance of taking

into consideration the grading practices they were using and which practices best

measured student achievement. However, the topic of what should be included in grades

remained a source of great contention (Leyrer, 2015). Teachers and experts had long

debated whether grades should be reflective of a student's performance in a multitude of

areas, including those such as behavior and effort, or whether grades should have been

reflective of a student's proficiency in a particular subject devoid of such non-academic

factors as behavior and effort ("Effective Grading Practices," 2011). Whereas some

teachers believed that non-academic factors should be included in grades to accurately

show all of a student's abilities, experts disagreed. Instead, experts recommended that teachers calculated grades on standards of achievement only (or, rather, on solely academic factors), separate from assessment of behavior, effort, and improvement (non-academic factors) (Brookhart, 2004; Guskey & Bailey, 2001; Marzano, 2010; O'Connor, 2009; O'Connor & Wormeli, 2011; Scriffiny, 2008; Wiggins & McTighe, 2006; Winger, 2005). In addition, experts emphasized that academic factors be separated from non-academic factors in grade reporting systems, as well (Brookhart, 2004; Guskey & Bailey, 2001; Marzano, 2010; O'Connor, 2009; O'Connor & Wormeli, 2011; Scriffiny, 2008; Wiggins & McTighe, 2006; Winger, 2005). O'Connor (2007) synthesized research and cited 15 best-grading practices aimed to support learning and correct inaccurate grading practices that resulted in distorted measurement of academic achievement, low-quality or poorly organized learning evidence, and inappropriate grade calculation.

Standards-based grading, a practice that was increasing in popularity in years recent to this writing, focused chiefly on students' proficiency on course objectives and graded only on achievement, yet still provided supplemental information about non-academic factors (Scriffiny, 2008). Guskey (2001) purported that standards-based grading systems, also known as criterion-referenced grading systems, most accurately and equitably reported student mastery, because in this kind of system, teachers usually assigned students two separate marks. Students received one mark indicating the teacher's assessment of the student's level of progress towards the learning standards (the academic factors) and received a second mark indicating the teacher's assessment of the student's level of progress towards the course expectations (the non-academic factors)

(Guskey, 2001). In this way, teachers separated academic and non-academic factors in grading, deliberately reporting them in a manner as to not combine the two.

Experts maintained that grading on achievement only, such as that occurring in a standards-based grading system, resulted in grades that more accurately communicated feedback about student mastery to all stakeholders, because it did not mix academic factors and non-academic factors that diminished what a grade should have indicated about a student's knowledge (Allen, 2005; O'Connor, 2007; Reeves, 2008; Wormeli, 2006). Therefore, experts considered grading and reporting systems that calculated grades based on clearly articulated standards of student learning as being the most effective (Guskey & Jung, 2010). This was because such grades and reporting systems altered the meaning of a grade from a singular, comprehensive assessment of student learning to a description of student performance on explicit skills (Guskey & Jung, 2010). Because it provided information that measured all students on comparable scales instead of providing a single letter grade that reflected student achievement on combined standards, experts found standards-based grading to be more objective than traditional grading (Yesbeck, 2011). For example, Scriffiny (2008) explained that standards-based grading enabled students to assume more ownership of their rights and responsibilities as a student and allowed for more discussion and reflection about learning.

According to experts, effective grading practices were those that resulted in grades based solely on students' mastery of the material in a specific subject, devoid of non-academic factors, such as student behavior and teacher expectations, in order to effectively communicate accurate, specific, meaningful feedback on what learning areas needed improvement (Allen, 2005; Marzano, 2000, 2007; O'Connor, 2007; Wormeli,

2006). Research supported that the most important purpose of grades was to provide

frequent, detailed feedback to students about specific objectives, standards, or learning

goals in relation to achievement (Brookhart, 2009; Guskey, 2004; Marzano, 2000;

McMillan, 2008; O'Conner, 2007; Wormeli, 2006). Guskey and Bailey (2001, 2010)

synthesized research and concluded that grading practice improvements were best

achieved through a comprehensive and multifaceted grade reporting system.

Therefore, experts recommended that if teachers feel that non-academic factors

were important to communicate to stakeholders to depict student learning, they should

have been reported separately from academic factors (Allen, 2005; Cross & Frary, 1999;

Guskey, 2006; Winger, 2005). This entailed the use of a supplemental grade or report by

which to communicate non-academic factors (Allen, 2005). This supplemental report

would have allowed teachers to communicate a student's achievement of non-academic

factors that were important to the learning process separately from academic factors, as

not to combine or confuse the two. Winger (2005) believed that communicating academic

and non-academic factors separately from one another would help protect the value of

grades and mitigated potential confusion surrounding their meaning, accuracy, and

perception.

The initial step to ensuring stakeholders that teachers' grading practices resulted

in fair and accurate grades was ensuring that teachers were using high-quality grade

reporting systems (Jung & Guskey, 2010). Jung and Guskey (2010) considered high-

quality grade reporting systems as those that possessed two important characteristics.

First, the grading system must have based grades off of clearly articulated standards of

student learning. Second, the grading system must have distinguished product criteria,

process criteria, and progress criteria. Experts agreed that these two components must have been satisfied and then reported separately (Guskey, 2006; Stiggins, 2007; Wiggins, 1996).

Guskey (2001) stated that, while standards-based grading and reporting systems were the best systems established to date to separate academic and non-academic factors in grades, there were challenges associated with using them. As a result, they were considered imperfect. One challenge when using standards-based grading was they were foreign to parents who were more familiar with the traditional grading system that used letter grades. As a result, most parents struggled to understand standards-based grading when first introduced to it and often said to teachers, "This is great. But tell me, how is my child doing really?" (Guskey, 2001, p. 4).

Parents' initial confusion could have been remedied by finding "a crucial balance in identifying standards that are specific enough to provide parents with useful, prescriptive information, but broad enough to allow for efficient communication between educators and parents" (Guskey, 2001, p. 8). Therefore, when using standards-based grading, it was imperative that stakeholders had a close working relationship in order for the information teachers were communicating to be accurately understood (Guskey, 2001). This required both that parents knew what the standards meant and how to interpret the levels of achievement or performance relative to the standards and that teachers ensured that parents were familiar with the language and terminology being used (Guskey, 2001). Guskey (2006) stated that, once they understood it, parents preferred separately reported factors, because it created a more comprehensive profile of their student's overall performance in school.

Similarly, employers and college admissions officers also preferred separately reported grades, because it distinguished students who earned high academic grades with relatively little effort in non-academic factors from those who earned equally high academic grades, while also satisfying non-academic factors (Guskey, 2006). Therefore, this standards-based grading helped protect teachers from claims against bias, because it helped ensure fairness and accuracy in grades (Guskey, 2006). Finally, Guskey (2006) maintained, reporting academic and non-academic factors separately would serve to bridge the existing gap between grades and state assessment scores by mitigating potential variation in what factors teachers were grading. Separately reported factors would strengthen the relationship between grades and scores, because there would be less room for teachers to blur meanings by mixing the academic and the non-academic. "Concerns about honesty and fairness compel us to reduce the mismatch between these two important measures of student knowledge and skill" (Guskey, 2006, p. 674).

Therefore, Guskey and Bailey (2001, 2010) emphasized that scores be reported by product, process, and progress criteria. More schools were adopting the practice of reporting separate grades or marks for these progress criteria (Guskey, 2006). This practice allowed teachers to isolate specific marks for academic factors and non-academic factors that pertained to learning; thereby, keeping assessments of achievement and performance distinct (Guskey, 2002; Stiggins, 2005). Guskey (2006) stated that such schools computed grade point averages and class ranks solely off of student academic

achievement or product grades. This resulted in "a better, more accurate, and much more comprehensive picture of what students accomplish in school" (Guskey, 2006, p. 673).

Another challenge associated with using standards-based grading and reporting systems was teachers' initial perception that it would create more work for them because the act of separating academic and non-academic factors in grading would have increased their work load (Guskey, 2001). However, standards-based grading required a considerable level of commitment, not additional work, from teachers in attempt to produce fair and accurate information about student mastery that could be communicated to stakeholders (Guskey, 2001). Guskey (2001) stated that standards-based grading systems did not lessen teachers' responsibility to assess student performance and report the results. Instead, standards-based grading systems posed the following unique challenges for teachers that must have been addressed to be most effective: (a) Changing from norm-referenced to criterion-referenced grading standards; (b) differentiating the kinds of grading criteria teachers used; (c) clearly specifying the purpose of each reporting tool used; (d) developing a report card that served as the cornerstone of a standards-based reporting system; and (e) determining the precise learning goals or standards on which to have based grades (Guskey, 2001).

Guskey (2006) found that teachers who used this procedure of separately reporting academic and non-academic factors in grades had actually found that it made grading easier and less work. He explained, when reporting academic and non-academic factors separately, teachers gathered the same evidence on student learning that they did when they calculated one overall grade. However, when reporting the factors separately, teachers no longer had to worry about the amount of weight given to a specific factor or

how to combine evidence (Guskey, 2006). Reporting these factors separately served to alleviate such burdens, thereby making grading easier for teachers (Guskey, 2006).

Guskey (2001) outlined four steps to be followed in developing standards-based grading: (a) Identification of the most important learning goals or standards that students would be expected to achieve at each grade level or in each course; (b) the establishment of performance indicators for the learning goals/standards; (c) the establishment of graduated levels of performance (benchmarks) for assessing each goal or standard; and (d) the determination of reporting systems that communicated teachers' judgments of students' progress in learning and culminating achievement relative to the learning goals or standards. Guskey (2001) maintained that the major advantages of a standards-based grading system were that it provided a clear description of student mastery and was useful for diagnostic and prescriptive purposes.

While standards-based grading was a system that could be employed by teachers of any content or grade level, it was most commonly used in elementary schools, because there was minimal content differentiation at this level (Guskey, 2001). Diversified curriculum at the middle and high school levels required standards-based reporting forms that varied from student to student, contingent upon their individual courses of study. Recent technological advancements, such as computer-generated reporting forms, allowed teachers to provide individualized reports, but relatively few middle and high school teachers used such (Guskey, 2001).

Experts recommended that teachers employed standards-based grading for exceptional learners, such as those possessing individualized education plans (IEPs), 504 plans, and English language learners (ELL) (Jung & Guskey, 2010). Teachers were able

to grade students who did not require special support or changes to the standards they were expected to achieve the same as they would any other student in the class with no change in the grading process. For students requiring special support or changes to the standards, teachers needed to determine what type of adaptation, consisting of either accommodation or modification, the standard needed. This was based on the circumstances of its use. If it was determined that accommodation was necessary, there was no change in the grading process required (Jung & Guskey, 2010).

If it was determined that modification was necessary, there were three steps the teacher should have completed (Jung & Guskey, 2010). First, the teacher should have determined the appropriate standard of what was believed the student could reasonably achieve by the conclusion of the academic year with special supports. Second, the teacher should have determined grades based on the modified standard, as opposed to the grade-level standard, to ensure appropriateness. Third, the teacher should have communicated the meaning of the grade in a manner that provided additional information to support modified standards and disclosed what was actually measured. Doing so clarified the meaning of the grade and fulfilled federal legislation requirements about reporting student performance on appropriately challenging learning tasks as well (Jung & Guskey, 2010).

However, despite expert recommendation that academic achievement should have been the only consideration in calculating a grade, a large discrepancy existed in reality (Allen, 2005; Brookhart, 1993, 1994; McMunn et al., 2003; O'Connor, 2007; Stiggins et al., 1989). According to Reeves (2008), teachers rarely used the best practices recommended by experts and supported by research. This was evidenced by the fact that,

while standards-based grading had been shown to be highly effective, it had not been

widely implemented (Reeves, 2008). The literature revealed that most teachers' grading

practices directly contradicted recommended practices (Webster, 2011).

For years, there were many expert recommendations as to how to improve grading

practices (Bloom, 1968; Brookhart, 2009; Guskey, 2006; Jongsma, 1991; McMillan,

2008; O'Connor, 2009; Tomlinson, 2005; Webster, 2011; Wiggins, 1994). In attempt to

assist schools in the development of fair-grading practices, the American Federation of

Teachers, the National Council on Measurement in Education, and the National

Education Association (1990) developed national professional standards for assessment -

two of which directly pertained to grading procedures (Webster, 2011). No matter the

source, expert-recommended grading practices had not been implemented, resulting in

the continued use of stagnant practices (Brookhart, 2009; Cizek et al., 1996; Stiggins,

2001). As stated by Webster (2011), "Teachers continue to use mostly traditional

practices" (p. 3).

Researchers purported possible reasons as to why this discrepancy between expert

recommendation and teacher practice existed: (a) Recommended practices may not have

reflected teacher opinions; (b) recommended practices may not have realistically

reflected classrooms; (c) teachers may not have been informed of the recommendations

and; (d) teachers may have lacked training in effective grading practices (Frary et al.,

1993; Friedman & Manley, 1991; Stiggins et al., 1989). As stated by Green and Emerson

(2007), "Grading is one of the least liked, least understood and least considered aspects of

teaching" (p. 495). As a result, teachers struggled with grading fairly (Brookhart, 2009)

and demonstrated concern over the topic of grading practices (Brookhart, 1994; Frary et al., 1993; Guskey & Bailey, 2001; Marzano, 2000).

Stiggins (1999) and Guskey (2006) plainly stated that teachers lacked adequate training to effectively grade students. This, coupled with the autonomous nature of current grading practices, left every teacher for themselves in determining how to grade (Guskey & Bailey, 2001; Stiggins, 2005; Truog & Friedman, 1996). Griswold (1993) maintained that teachers may have disregarded expert recommendations, because they lacked input. Consequently, he suggested that experts should have sought input from teachers and made any necessary adjustments in attempt to get teachers to embrace desired practices.

Heifetz and Linsky (2002) agreed with Griswold (1993) that teachers who took issue with, or failed to follow, recommended grading practices were problematic and this represented one of the major issues about current high school grading practices. They emphasized the following obstacles: (a) Recommendations and practices were incompatible; (b) changing practices demanded difficult learning; (c) people who represented the problem could also have represented the solution and; (d) the act of changing practices could have resulted in further disconnect. However, under the influence of informed school leaders, actions could have been taken to gain support for adaptive change (Heifetz, Grashow, and Linsky, 2009). Heifetz et al. (2009) further stated that challenge and understanding were crucial components of the reform process, because they helped leaders identify teachers' reasons for grading the way that they did, in order to determine the most efficient way to correct their practices.

In attempt to remedy the discrepancy between recommendation and practice, Allen (2005) suggested teachers receive professional development about effective grading practices. Professional development facilitated the growth of teachers' expertise and informed them about best practices (Yesbeck, 2011). Allen (2005) explained that professional development could have helped teachers to identify and create effective assessments, by which to accurately and objectively measure student achievement. The literature supported professional development as a medium by which teachers could have discussed, analyzed, and aligned grading practices to best measure student achievement (Yesbeck, 2011). However, as noted by Yesbeck (2011), professional development for teachers did not begin once teachers were hired to teach, it began in teacher preparation programs and should have continued throughout the course of their careers.

Gender

The first teachers in America were women, because the position was regarded as a female occupation, due to its focus of working closely with children, especially at the elementary level (National Women's History Museum, 2007). The role of teacher, outside of college and university teaching, was entirely reserved for unmarried, native-born women, who started working in their late teens and early twenties before getting married and leaving the work force (National Women's History Museum, 2007). At the same time, men were raised to consider teaching a job for women and were taught that it was more honorable for men to work in labor-related or white-collar jobs (Patrick, 2009). As a result of this prominent ideology, teaching became a female-dominated profession and it continued to remain such (Patrick, 2009). Feistritzer (2011, p. 15) further pointed

out that not only was the profession female-dominated, but it was dominated by mostly Caucasian females (p. x).

Although, in decades recent to this writing, several professions that were traditionally male-dominated have expanded to include more females, this has not changed the longstanding fact that, in the teaching profession, most teachers were women (Ingersoll & Merrill, 2010). Patrick (2009) posited that this overwhelming disparity in the amount of male teachers presented a substantial disadvantage for male students, in particular due to the minimal amount of male role models it provided for them. Ingersoll and Merrill (2010) expanded upon Patrick, suggesting that the overall lack of male teachers encountered by students throughout their education deprived disadvantaged students of surrogate fathers.

In attempt to have better understood the composition of the teaching workforce, the NCES conducted six cycles of the Schools and Staffing Survey (SASS) throughout 1987 to 2008 and collected teacher demographic data for analysis. According to 2007-2008 data, the number of female teachers increased from 66% in 1980 to 76% in 2007-2008 (as cited in Ingersoll & Merrill, 2010). The data also revealed that the secondary level of education had seen the greatest increase in the number of female teachers, replacing male teachers who constituted the majority until the late 1970s (Ingersoll & Merrill, 2010).

In comparison, there were only minimal increases in the amount of female teachers employed at the elementary level, where they had historically comprised the majority of teachers (Ingersoll & Merrill, 2010). Ingersoll and Merrill (2010) stated that the reason why the teaching profession was experiencing an increased level of female

domination was unclear. However, they attributed the growth to three contributing

factors: (a) an increase in the total number of females entering the workforce overall; (b)

the profession's shortened workday and seasonal breaks that offered a more manageable

career for mothers with young children; and (c) the increase in career opportunities and

promotions available to females, as evidenced by the considerable increases in the

number of female administrators.

Regardless of their gender, most teachers possessed gender biases believed to

have developed from cultural norms that often manifested in the classroom (Scantlebury,

2009). Gender bias occurred "when people make assumptions regarding behaviors,

abilities or preferences of others based upon their gender" (Scantlebury, 2009, p. 1). For

example, teachers' gender role stereotypes of students provided that males were

boisterous, academically able, yet socially uncommunicative, whereas females were

quiet, polite, studious, and possessed greater social skills (Scantlebury, 2009). Most

teachers expected male students to excel at math and science and expected female

students to excel at reading and language arts. Therefore, teachers' acceptance of specific

behaviors from one student or another was based upon the student's gender. This was

why female students who presented discipline issues for teachers and male students who

were quiet and studious often experienced a lack of understanding from teachers

(Scantlebury, 2009).

Researchers found that a student's achievement was often dependent on the

gender of the student's teacher (Dee, 2006). According to cognitive scientists, teachers

may subtly have communicated differing academic expectations of male and female

students that helped shape the achievement gap wherein boys scored higher in math and

science and girls scored higher in reading. Dee (2006) purported that a teacher's gender had significant effects on a student's performance on tests, teacher perception of students, and students' engagement with academic work. When taught by a teacher of the same gender, students' engagement increased, they behaved more appropriately, and performed at a higher level. These findings persisted, even after accounting for a plethora of other characteristics that may have influenced student learning (Dee, 2006). Therefore, the teacher's gender established biased expectations for students that impacted their academic achievement, and by extension, their grade.

This was further complicated by the facts that most teachers were female and the number of male teachers had been consistently declining throughout the past several decades, at the time of this writing (Patrick, 2009). According to 2011-2012 research conducted by the U.S. Department of Education's NCES, there were about 3.4 million public school teachers in the U.S (as cited in Feistritzer, 2011, p. vii). According to Feistritzer (2011, p. 12), 84% of public school teachers were female (p. xi). In 2012, the U.S. Bureau of Labor found that male public school educators represented only 2.3% of pre-kindergarten and kindergarten teachers, 18.3% of elementary and middle school teachers, and 42% of high school teachers (p. 3).

This was important information considering that when children began kindergarten, the two genders performed similarly on tests of both mathematics and reading (Dee, 2006). However, by the time students reached their senior year of high school, the gender gaps in achievement had grown significantly. In reading, 17-year-old males scored 31% of a standard deviation below 17-year-old females, a deficit equal to approximately one grade level (Dee, 2006, p. 70). In math and science, females of that

age scored 22% and 10% of a standard deviation lower than males (Dee, 2006, p. 70).

Dee (2006) posited that "if half of the English teachers in grades 6, 7, and 8 were male

and their effects on learning were additive, the achievement gap in reading would fall by

approximately a third by the end of middle school" (p. 75).

When taught by a female, boys were more likely to be viewed as disruptive,

whereas girls were less likely to be seen as either disruptive or inattentive (Dee, 2006).

Boys also had fewer positive reactions toward the relative academic subject when there

was a female teacher and reported that they did not look forward to the class. Similarly,

when being taught by a male, girls were more likely to state that they did not look

forward to a subject, that it was not useful for the future, or that they were scared to ask

questions (Dee, 2006).

In attempt to mitigate the achievement gap, some educational advocates had

suggested that schools employ single-sex classrooms, wherein the teacher and pupils

shared the same gender orientation. However, Dee (2006) warned that single-sex learning

environments may have adverse consequences regardless of any potential positive

impacts. This was largely due to the fact that gender dynamics could differ, even in a

single-sex classroom. As part of a reform movement to combat gender biases in teacher

behavior and expectations, Dee (2006) presented the idea that teachers underwent gender-

specific training based on evidence concerning the varying learning styles of male and

female students. As stated by Scantlebury (2009), "Teachers are critical components in

challenging gender bias in schooling" (p. 2).

Age

During the 1987-1988 school year, the age distribution of public school teachers resembled a tall peak with the modal age of 41; but, by 2007-2008 the modal age had increased to 55 (Ingersoll & Merrill, 2010). Since about 2005, there was a trend of younger teachers aged 30 or younger entering the profession that were replacing the older, retiring teaching force (Feistritzer, 2011). As a result, the amount of teachers that were 30-years-old or younger rose significantly from 2005 to 2011, resulting in one in five teachers in 2011 being under the age of 30 (Feistritzer, 2011, p. 12). In 2012, the NCES reported that 44% of public school teachers were under the age of 40 (National Center for Education Statistics [NCES], 2019). This meant that many retiring teachers were being replaced by teachers in their 20s and 30s with relatively little experience. Therefore, in addition to an aging teacher workforce, America was also experiencing a greening teaching profession (Ingersoll & Merrill, 2010).

However, the teaching force was growing at a remarkable rate, due to the amount of beginning teachers that entered the profession in recent years. In 1987-1988, the modal teacher possessed 15 years of teaching experience (Ingersoll & Merrill, 2010). Whereas by 2007-2008, the modal teacher was beginning his or her first year of teaching (Ingersoll & Merrill, 2010).

Educational Level

Overall, the public school teaching force became more educated in terms of the highest degree held. In 2005, 47% of public school teachers possessed a master's degree as the highest degree earned (Feistritzer, 2011, p. 19). In 2011-2012, the NCES found this number had increased, reporting that 56% of teachers possessed a master's degree or

higher (NCES, 2019). While more teachers had been earning higher degrees, the means

by which some teachers were obtaining their certification had changed throughout recent

years.

In order to teach in public K-12 schools in the United States, one was required to

possess a bachelor's degree and a teaching certificate in the state in which one was

teaching. The traditional means of obtaining certification to teach consisted of earning a

bachelor's degree in education (Feistritzer, 2011). However, alternative routes to teacher

certification began to appear in the mid-1980s and had taken off by the late 1990s

(Feistritzer, 2011). These alternative routes consisted of state-defined routes, through

which an individual who already possessed at least a master's degree could obtain

certification to teach without having to return to college and complete a campus-based

teacher education program (Feistritzer, 2011).

The number of teachers graduating from these alternative programs had steadily

increased, beginning around the turn of the millennium, to 16% as of 2011 (Feistritzer,

2011, p. 21). In 2011, the National Center for Education Information (NCEI) reported

that 95% of teachers with 25 or more years of experience had earned their teacher

certification through a traditional campus-based undergraduate (82%) or graduate (13%)

program (as cited in Feistritzer, 2011, p. 22). Conversely, four out of 10 teachers with

five or less years of experience entered the teaching force after completing alternative

programs (Feistritzer, 2011, p. 22).

Interestingly, more men than women were entering teaching through alternative

routes. According to 2011 data obtained by the NCEI, 32% of men, compared to 22% of

women, entered the teaching force through an alternative route (as cited in Feistritzer,

2011, p. 23). Furthermore, 61% of teachers who obtained their certification through an alternative route hold a bachelor's degree in a field other than education as their highest degree earned (Feistritzer, 2011, p. 29).

Overall, there was a considerable range in the number of college semester hours of education courses teachers had taken (Feistitzer, 2011). 2011 NCEI data showed that 40% of all surveyed public school teachers reported that they had taken 50 or more college semester hours of education courses, an additional 31% reported that they could not remember how many hours of college education courses they had taken, and 11% reported that they had taken fewer than 25 college semester hours of education courses (as cited in Feistritzer, 2011, p. 32). Of teachers who earned their teacher certification through an alternative program, 20% reported that they had taken 50 or more hours of education courses, 44% reported that they had taken fewer than 25 hours of education courses, and 12% reported that they had not taken even a single education course (Feistritzer, 2011, p. 32).

Feistritzer (2011) stated that given this range in the number of college semester hours of education courses taken by public school teachers, it seemed appropriate that teachers ranked such courses, and the faculty that taught them, as particularly low in terms of how valuable they were in developing pre-service teachers' competencies to teach. However, while the exact relationship between teachers' academic ability and teaching quality was unclear, academic ability was generally regarded to be an important indicator of the competence of employees in many lines of work, including teaching (Ingersoll & Merrill, 2010). In addition, the caliber and selectiveness of the institution from which teachers completed their education was also generally regarded as indicative

of an individual's academic ability (Ingersoll & Merrill, 2010). Ingersoll and Merrill (2010) found that there were distinctions by gender in regards to the academic abilities of first-year teachers, citing an overall decrease in the academic abilities of first-year male teachers from 1987-2008 and an overall increase in the academic abilities of first-year female teachers throughout this time.

Years of Service

As revealed by the SASS administered in 2007-2008 (the latest year for which these data were published), approximately 146,500 of teachers were new hires with no prior teaching experience, 92,500 were new college graduates, and 54,000 fell under the category of "delayed entrants" meaning that they had earned a degree but had not started teaching straight out of college (Feistritzer, 2011, p. viii). The amount of public school teachers possessing five or less years of teaching experience increased from 18% in 2005 to 26% in 2011, while the amount of teachers possessing 25 or more years of experience dropped from 27% in 2005 to 17% in 2011 (Feistritzer, 2011, p. 19). In addition, the routes by which non-veteran teachers were obtaining their certification differed from that of veteran teachers with the increased number of alternative certification routes available to non-veteran teachers (Feistritzer, 2011).

This was important information, because non-veteran teachers (those who possessed less than 10 years of full-time teaching experience in a public school setting), were considerably more open to proposed grading reforms than were veteran teachers (those who possessed 10 or more years of full-time teaching experience in a public school setting) (Feistritzer, 2011). Hargreaves and Fullan (2013) maintained that teachers who possessed one to three years of teaching experience were more enthusiastic than at any

other juncture in their career. "They are more committed, more dedicated" (Hargreaves & Fullan, 2013, p. 38). However, on average, they were less competent. Also, on average, around the 22-years-of-service milestone, teachers' commitment was declining and their capabilities varied greatly (Hargreaves & Fullan, 2013).

Furthermore, increasing teacher turnover was placing more non-veteran teachers in the classroom. Ingersoll and Merill (2010) stated that teacher turnover, which included teachers moving from one school to another and teachers who left the profession altogether, had been increasing since the early 1990s. While average teacher turnover rates fluctuated from year-to-year and from state-to-state, as of 2005 they had increased overall by 28% (Ingersoll & Merill, 2010, p. 18).

Content Area

Researchers found that approximately half of public school teachers graduating from traditional college teacher education programs taught elementary school (Feistritzer, 2011). Conversely, teachers that earned their credentials through alternate programs were more likely to teach high school in the subjects considered high demand - math, science, special education, and bilingual education (Feistritzer, 2011). Despite the orientation of their teacher education program, overall, the percentage of teachers who were females remained high in the elementary and middle school levels (Dee, 2006). In high school, the percentage of teachers who were female dropped and teachers were more likely to be men, especially in science and history classes (Dee, 2006). This was despite the fact that the increase in female teachers had been concentrated at the secondary level (Ingersoll & Merrill, 2010).

Male and female students' participation in and perceptions of core content area subjects (math, science, English, and social studies) varied in ways that paralleled gender gaps in academic achievement (Dee, 2006). This distinction was strongest in science, where students reported that female science teachers were substantially more effective in advocating girls' engagement in this field. Female students were more likely than male students to report they were scared to ask questions in science, as well as math and social studies. In addition, they were less likely than male students to look forward to these classes or to perceive them as being useful for their future. On the other hand, male students were more likely than female students to report more negative perceptions of English classes (Dee, 2006).

Scantlebury (2009) maintained that in math and science classes, teachers often demonstrated gender bias toward students by promulgating a myth that males were naturally better at math and science than females. The implication of doing such was, if girls achieved academic success in these subjects, it was the result of hard work, not their intelligence, whereas boys' achievement was accredited to their natural talent (Scantlebury, 2009). While there had been some evidence that gender bias in these subjects might have been decreasing due to the increased percentage of female students participating in biology, chemistry and algebra, subjects that served as prerequisites for college majors, such as engineering and physics, continued to be dominated by male students (Scantlebury, 2009).

Collaboration

In 2011, NCEI teacher survey respondents were asked to rate the effectiveness of 15 aspects of their teacher preparation programs. The surveys revealed that "Discussions

with fellow teachers" was regarded by the respondents as most important in preparing them to be effective teachers, followed by the "actual teaching part of the program" (as cited in Feistritzer, 2011, p. 30). Similarly, the NCEI (2011) also found that 76% percent of the teachers surveyed ranked working with other teachers and colleagues as "very valuable" in developing competence to teach second only to their own teaching experiences (as cited in Feistritzer, 2011, p. 30).

Research supported this sentiment regarding the importance of teacher collaboration and further showed that the most essential ingredient in school reform was collaborative time for teachers during which they could have undertaken and sustained school improvement (Raywid, 1993). During collaborative time, teachers were able to participate in discussions or book talks related to current educational topics, such as grading practices (Yesbeck, 2011). Over time, these discussions had the potential to develop into grading reform initiatives that could have started within one department and spread to other departments or the whole school (Yesbeck, 2011).

Professional Learning Communities (PLCs) and professional development were two types of teacher collaboration that provided time for teacher reflection - an important component of teaching (Yesbeck, 2011). Yesbeck (2011) stated that this reflection should have focused on "1. Why teachers grade, 2. What student grades represent in terms of student achievement, and 3. How they define the primary purpose of grades" (pp. 129-130). After reflecting, teachers should have considered the academic and non-academic factors they included in their grading practices and compared and contrasted such with measurement experts' recommendations, grading policies, and other teachers, as the act of doing so may have resulted in teachers' discovery that their current practices

inaccurately represented student achievement (Yesbeck, 2011).

Raywid (1993) synthesized literature and stated, "Collaborative time for teachers to undertake and sustain school improvement may be more important than equipment or facilities or even staff development" (p. 30). As stated by Davis (2015), as more rigorous learning standards were adopted throughout the nation, schools' dependency on teachers to raise student achievement increased. However, this was problematic given that many schools failed to provide teachers with more than just a few minutes a day in which to engage in collaborative work (Davis, 2015).

Raywid (1993) stated that in order for school reform to be successful, including in the area of grading practices, "teachers collectively must be involved in its implementation (p. 30). However, even in times when change was not actively occurring, teacher collaboration time was still imperative (Raywid, 1993). Many (2009) stated, "One of the critical conditions for the development of collaborative cultures is designated and protected time for teachers to meet and collaborate during the regular school day" (p. 8). Therefore, successful schools could have been distinguished from unsuccessful schools by the frequency and extent to which teachers discussed practice, collaborated to design materials, and informed and critiqued each other's practices (Little, 1982). While some schools had simply added meeting days to their annual school calendar, others increased the amount of teachers they employed and developed a rotation schedule for more teacher collaboration time to occur throughout the school day (Raywid, 1993).

Faced with the dilemma of providing more collaboration time without considerably increasing school costs, schools throughout the nation were experimenting with innovative ways in which to make or find time for teacher collaboration (Raywid,

1993). However, experts stated that time would not merely have been found to collaborate - it must have been created (Many, 2009). As a result, recent years had seen the exponential growth of professional learning communities (PLCs) implemented at schools (Panagos, 2011).

However, schools were expected to meet specific criteria in order to be considered a PLC. Hord (1997) maintained that the community must have been engaged in a multitude of activities "including sharing a vision, working and learning collaboratively, visiting and observing other classrooms, and participating in shared decision making" (p. 1). PLCs focused on the learning of each student (DuFour, DuFour, Eaker, & Many, 2006). They offered a common vision for educators; faculty at schools working in a PLC shared an understanding and commitment to school improvement (Lunenburg, 2010). These qualities, coupled with their positive effect in promoting school reform, had expanded the popularity of PLCs (Hord & Sommers, 2008).

Lunenburg (2010) stated that research on improving schools "boils down to the ability of the people within a school to function as a PLC" (p. 1). Schools with the strongest PLCs, as measured by teachers' attitudes toward their peers and individual development as educators, consistently generated higher student performance (Davis, 2015). Given this information, it was not surprising that some people considered PLCs to be the most effective institution for school development concentrated on student achievement (Panagos, 2011). This sentiment was reflected in a 2015 survey of state teachers of the year that identified collaboration as one of the top three areas they believed that school funding should be focused on in attempt to promote student achievement (Davis, 2015).

While PLCs were a popular method of collaboration, it was important for schools to remember there was more than one way to collaborate (Hargreaves and Fullan, 2013). No matter the form of collaboration, Hargreaves and Fullan (2013) emphasized the importance of professional capital, which could have been gained through the process of teacher collaboration and its association with positive outcomes in student achievement. Hargreaves and Fullan (2013) defined professional capital as the sum of three different kinds of capital: "1. human capital (the talent of individuals such as their qualities; qualifications and competencies on paper), 2. social capital (the collaborative power of the group); and 3. decisional capital (the wisdom and expertise to [have made] sound judgments about learners that [we]re cultivated over many years)" (Hargreaves & Fullan, 2013, p. 37). Their findings built off of research conducted by Leana (2011) which indicated a relationship between human and social capital.

Leana (2011) found that schools with high social and human capital experienced higher student achievement outcomes in comparison to schools with lower social and human capital. Leana (2011) also found teachers with low human capital who were working in schools with higher social capital experienced better outcomes than teachers working in schools with lower social capital. Therefore, "being in a school around others who are working effectively rubs off on teachers and engages them" (Hargreaves & Fullan, 2013, p. 37). This type of environment was found to be conducive to school reform because it encouraged teachers to become agents of change (Yesbeck, 2011).

Conclusion

This review of history and literature of grading practices and the demographic characteristics of teachers explained the need for school reform and sustained improvement in teachers' grading practices. By briefing the reader on measurement

experts' grading practice recommendations and models for collaboration, the researcher hoped to provide the reader with a foundation for school reform. This foundation may assist educational professionals in understanding the need for sustained improvement in grading practices and the importance of reform and teacher acceptance regarding such. This study was an effective way in which to view teachers' grading practices and expert recommendations from the teachers' perspective. It specifically identified the academic and non-academic factors teachers considered in their grading practices, demographic characteristics of teachers, teachers' perceptions of expert-recommended grading practices, and the amount of time and support districts offered in relation to implementation of best practices.

Chapter Three: Methodology

Overview

Experts recommended grading practices emerged as a result of teachers' varying grading practices, wherein academic and non-academic factors were often combined into a single grade. The review of the literature demonstrated both a need for school reform in grading practices and a notion of reluctance by teachers to change their then-current grading practices. However, teacher leaders could have initiated reform efforts (Hills, 1991) wherein teachers served as agents of change (Yesbeck, 2011).

This mixed methods study of grade 7th through 12th core content area teachers at four Missouri school districts focused on obtaining the demographic characteristics and insights of teachers, as well as information about their then-current grading practices, to contribute to the reform process. The researcher cross-referenced then-current teacher grading practices to the level of implementation of best-grading practices, as it pertained to the amount of time a district provided for teacher collaboration. Multiple content areas and the range of 7th through 12th-grade levels were chosen for the purpose of providing a depth of range with meaningful and diversified perspectives of which to compare and contrast, as opposed to studying only one content area and/or grade level.

Prior to collecting data, Instructional Review Board (IRB) approval was obtained from Lindenwood University (see Appendix J). After gaining IRB approval, the researcher contacted each of the districts' superintendents to seek their participation in the study (see Appendix I). Permission was granted from each of the four districts.

Grading practices were exercised by all teachers. All member schools belonging to SE RPDC emphasized effective teaching/learning practices. Therefore, all grade 7th

through 12th core content area teachers in the four selected school districts were invited
to participate in the study.

Participants were informed about the study through the Model Recruitment
Statement (see Appendix A). This was contained in an email that was sent to them from
their superintendent. According to Creswell (1998), discovering participants who were
experiencing the phenomenon being studied and who were easily assessable was
consistent with the phenomenological approach. The researcher selected interview
participants by contacting teachers recommended by their respective superintendents.

The superintendents forwarded the link to the researcher's online survey to their
district teachers via email. Teachers were then selected by their respective superintendent
to complete qualitative interviews. Participation was voluntary.

Chapter Three presents the methodology of the study, and provides information
on the research design, descriptions of the participants, the selection process, and the
setting. Strategies for analysis and synthesis of data and limitations of the study are also
detailed.

The Research Site

Four school districts in the 16 counties that were members of Missouri's SE
RPDC were selected as the research sites for this study. These schools focused on school
reform in the form of a Collaborative Work (CW) Project, wherein they emphasized
common formative assessment, data-based decision making, and effective teaching/
learning practices. The teachers within the districts taught at schools considered either
rural or suburban.

The researcher gave each of the schools a pseudonym, which were the names used throughout this study: Washington, Adams, Jefferson, and Madison. The researcher worked closely with teachers at these sites to gain their perspectives and insights regarding the topic of grading practices. McMillan and Schumacher (2006) found that site selection was particularly important when conducting qualitative interviews. Therefore, the qualitative interviews in this study were conducted in a quiet setting that limited interruptions and noise.

It was the hope of the researcher that the data obtained from this study would be used to improve grading practice reform efforts at these schools and any other schools that sought to undertake such action. In addition, the researcher hoped the data obtained would be used to guide professional development about grading practices. Finally, the researcher hoped the data obtained would provide insight into the gap that existed between actual and expert-recommended grading practices.

Developing the Intervention

According to Bogdan and Biklen (2007), qualitative research possessed specific traits, including: (a) a naturalistic element (understanding that resulted from immersion in the natural environment; (b) descriptive data; (c) emphasis on the process; (d) is inductive; and (e) has meaning. Muijs (2010) found that non-numerical data also played a critical role in qualitative research. Therefore, in addition to the online survey, in this study, the researcher employed a phenomenological design and interviewed six teachers. This was completed in an attempt to better describe and analyze the teachers' reasons for grading as they did. As stated by Panagos (2011), "Qualitative researchers . . . see how people make sense of their lives, looking for the participants' perspectives" (p. 38).

The phenomenological design of this study pertained to the descriptions and analysis of teachers' perceptions of grading practices and the support and services a district should have provided to facilitate implementation of best-grading practices. McMillan and Schumacher (2006) stated that a phenomenological design was appropriate when concentrating on a singular, shared phenomenon in order to have gained understanding. The teachers in this study satisfied this criterion, because they all calculated and reported student grades.

In contrast to qualitative research, Aliaga and Gunderson (2002) described quantitative research as explaining a phenomenon by collecting numerical data that were analyzed using mathematically-based methods, such as statistics. Quantitative methods were considered particularly useful when the researcher was seeking both "depth and meaning" (Muijs, 2010, p. 11). Quantitative methods were commonly combined with qualitative methods in social science research, because using one or the other method on its own would have inadequately addressed the complex problems explored; their combination resulted in an expanded understanding of research problems (Creswell, 2009).

This research concentrated on understanding teacher grading practices, gaining insight into the gap that existed between expert recommendations and then-current grading practices, and discovering which resources districts should have provided to facilitate the implementation of best-grading practices. Therefore, the researcher used both qualitative and quantitative data to answer the hypotheses and research questions. Doing so allowed the researcher to provide a more comprehensive description of the

context. As a result, this study possessed degrees of both qualitative and quantitative inquiry.

According to Creswell and Plano Clark (2007), mixed methods research design was a flexible method developed under the assumption that the combining of qualitative and quantitative approaches provided a better depiction of research problems than either approach was able to do when considered individually. Miles and Huberman (1994) stated, when the two types of data were combined, the result was "a powerful mix" (p. 42). This mix would have simultaneously preserved the study and enabled the researcher to have described the phenomenon in question (Panagos, 2011). Therefore, Muijs (2010) contended, "when we want to look at both breadth and depth, or at both causality and meaning . . . it is best to use a so-called mixed methods design in which we use both quantitative and qualitative methods" (p. 9).

Creswell (2002) further explained that the purpose of having used a mixed methods design was to "simultaneously collect both qualitative and quantitative data, merge the data, and use the results to best understand a research problem" (p. 564-565). Panagos (2011) summarized, "The purpose in mixing methods is to obtain enough data to accurately tell the story" (p. 40). And, as stated by Barth (2003),

> People tell stories about events that have left an impression on their lives. By listening, one places value in the experiences of another. Craft knowledge is the collection of wisdom and insights one accumulates by showing up on the job. If ways can be found to unlock, celebrate, and exchange craft knowledge, how much better each of us can perform our work. Storytelling is one way. Every story- and every storyteller - has value. (pp. 2-3)

Therefore, this research was designed to enable the teachers to tell the researcher their stories of their grading practices in attempt to gain a better understanding of the phenomenon that was grading.

Research Questions and Null Hypotheses

Null H1: There will be no differences in demographic characteristics within respondents to a Likert-scale survey, with regard to best-grading practices, among the following teacher-related variables: Teachers' age, teachers' gender, teachers' levels of education, teacher's years of service, and teachers' content areas.

Null H2: There will be no differences in perception of best-grading practices, measured by a Likert-scale survey, among the following support structures: Frequency of teachers' discussions with peers regarding best-grading practices, frequency of teachers' discussions with building administrators regarding best practices, and frequency of the teachers' engagement in self-reflections in which the teacher is engaged.

Null H3: There will be no differences in perception of agreement topic-to-topic, and agreement-to-disagreement with survey statements, with regard to best-grading practices, measured by a Likert-scale, with regard to the following support characteristics: A student's behavior, grades as a means of punishment, student responsibility, a student's ability level, assignment completion points, extra credit, a teacher's personal judgment, assignments other than assessments, a student's involvement in extracurricular activities, zeroes for missing assignments, a student's GPA, class participation, a student's gender, a student's physical appearance, homework, note-taking, aesthetic appearance of a student's work, a student's personality, student effort, using grades as a reward, student progress towards achieving the course learning

standards, individual grades for group work, individual projects, project-based grading, project-based assignments, student portfolios, conferences with students about individual grades throughout each grading period, student-set personal learning goals for the course, student personal reflection on achievement of the course learning standards, district-provided time for teachers in the same content area to meet to align grading practices, district-provided professional development about best-grading practices, pre-service training for teachers on best-grading practices, the belief that teachers should be accountable for following best-grading practices, teacher-to-teacher grading practice comparisons, district-provided common planning time for teachers in the same content area, district-provided collaboration time for teachers on early release days, district-provided collaboration time for teachers on late start days, district-provided collaboration time for teachers on professional development days, teacher collaboration, the belief that grades should be a direct reflection of student mastery of course learning standards, grades as a means to provide feedback to stakeholders, gradebook field that indicates student mastery of learning competencies, re-do assignments, test re-takes, and summative and formative assessments.

Null H4: There will be no relationships among levels of agreement topic-to-topic, with regard to best-grading practices, measured by a Likert-scale survey, with regard to the following support characteristics: A student's behavior, grades as a means of punishment, student responsibility, a student's ability level, assignment completion points, extra credit, a teacher's personal judgment, assignments other than assessments, a student's involvement in extracurricular activities, zeroes for missing assignments, a student's GPA, class participation, a student's gender, a student's physical appearance,

homework, note-taking, the aesthetic appearance of a student's work, a student's

personality, student effort, using grades as a reward, student progress towards achieving

the course learning standards, individual grades for group work, individual projects,

project-based grading, project-based assignments, student portfolios, conferences with

students about individual grades throughout each grading period, student-set personal

learning goals for the course, student personal reflection on achievement of the course

learning standards, district-provided time for teachers in the same content area to meet to

align grading practices, district-provided professional development about best-grading

practices, pre-service training for teachers on best-grading practices, the belief that

teachers should be accountable for following best-grading practices, teacher-to-teacher

grading practice comparisons, district-provided common planning time for teachers in the

same content area, district-provided collaboration time for teachers on early release days,

district-provided collaboration time for teachers on late start days, district-provided

collaboration time for teachers on professional development days, teacher collaboration,

the belief that grades should be a direct reflection of student mastery of course learning

standards, grades as a means to provide feedback to stakeholders, gradebook field that

indicates student mastery of learning competencies, re-do assignments, test re-takes, and

summative and formative assessments.

  RQ1: What types of support work best for teachers trying to implement best-

grading practices?

  RQ2: Which academic factors do teachers use when grading?

  RQ3: Why do teachers choose to use academic factors when grading?

  RQ4: Which non-academic factors do teachers use when grading?

RQ5: Why do teachers choose to use non-academic factors when grading?

Data Collection and Analysis Procedures

The researcher worked closely with the teachers to uphold two of Patton's (2008) concepts of "establish[ing] trust and rapport" while "maintain[ing] distance to guarantee objectivity and credibility" (p. 36). No personal identification information, other than subgroup demographics, were collected about the participants. All data will have been stored in a secure location for three years and will have been destroyed after said time. Data were reported in an ethical manner with minimal or no risk incurred by the participants. All teachers in the study gave their consent to participate (see Appendices A and B). It was expected that the teachers provided honest answers to the questions in the study.

The data collection consisted of two phases. First, was the online survey. Second, were the interviews. Participants were in no way rushed to answer survey or interview questions. Making the survey available via the internet allowed for participants to complete the questions online at their convenience. Interviews were scheduled according to the interviewee's preference to accommodate their time constraints and availability.

Teachers completed the online survey that generated data points for the research via *Qualtrics*. They were given a link to the survey instrument with written directions and asked to complete the data electronically, so that they were more easily disaggregated. The online survey consisted of 60 questions. The responses for the majority of the question prompts used a Likert scale wherein a response of 1 indicated strong disagreement and 5 indicated strong agreement.

Qualitative, in-depth interviews occurred via telephone. Participants answered questions from an interview guide during the audio-taped interviews (see Appendix F). The questions were crafted to be open-ended which enabled both the participants to provide more in-depth responses and the researcher to ask more probing questions. The order in which the interview questions were asked varied to promote conversational flow (Creswell, 2007). Once all six interviews were concluded, the researcher coded and analyzed the data for common themes.

The study focused on four hypotheses and five research questions. They were constructed to describe the participants' real-life experiences of the phenomenon (Creswell, 1998). The research questions and hypotheses are presented below.

Null H1: There are no differences in demographic characteristics within respondents to a Likert-scale survey, with regard to best-grading practices, among the following teacher-related variables: Teacher's age, teacher's gender, teacher's level of education, teacher's years of service, and teacher's content area.

Null H2: There are no differences in perception of best-grading practices, measured by a Likert-scale survey, among the following support structures: Frequency of teacher's discussions with peers regarding best-grading practices, frequency of teacher's discussions with building administrators regarding best practices, and frequency of the teacher's engagement in self-reflection in which the teacher is engaged.

Null H3: There are no differences in perception of agreement topic-to-topic, with regard to best-grading practices, measured by a Likert-scale survey, with regard to the following support characteristics: A student's behavior, grades as a means of punishment, student responsibility, a student's ability level, assignment completion points, extra

credit, a teacher's personal judgment, assignments other than assessments, a student's involvement in extracurricular activities, zeroes for missing assignments, a student's GPA, class participation, a student's gender, a student's physical appearance, homework, note-taking, aesthetic appearance of a student's work, a student's personality, student effort, using grades as a reward, student progress towards achieving the course learning standards, individual grades for group work, individual projects, project-based grading, project-based assignments, student portfolios, conferences with students about individual grades throughout each grading period, student-set personal learning goals for the course, student personal reflection on achievement of the course learning standards, district-provided time for teachers in the same content area to meet to align grading practices, district-provided professional development about best-grading practices, pre-service training for teachers on best-grading practices, the belief that teachers should be accountable for following best-grading practices, teacher-to-teacher grading practice comparisons, district-provided common planning time for teachers in the same content area, district-provided collaboration time for teachers on early release days, district-provided collaboration time for teachers on late start days, district-provided collaboration time for teachers on professional development days, teacher collaboration, the belief that grades should be a direct reflection of student mastery of course learning standards, grades as a means to provide feedback to stakeholders, gradebook field that indicates student mastery of learning competencies, re-do assignments, test re-takes, and summative and formative assessments.

Null H4: There are no relationships among levels of agreement topic-to-topic, with regard to best-grading practices, measured by a Likert-scale survey, with regard to

the following support characteristics: A student's behavior, grades as a means of

punishment, student responsibility, a student's ability level, assignment completion

points, extra credit, a teacher's personal judgment, assignments other than assessments, a

student's involvement in extracurricular activities, zeroes for missing assignments, a

student's GPA, class participation, a student's gender, a student's physical appearance,

homework, note-taking, the aesthetic appearance of a student's work, a student's

personality, student effort, using grades as a reward, student progress towards achieving

the course learning standards, individual grades for group work, individual projects,

project-based grading, project-based assignments, student portfolios, conferences with

students about individual grades throughout each grading period, student-set personal

learning goals for the course, student personal reflection on achievement of the course

learning standards, district-provided time for teachers in the same content area to meet to

align grading practices, district-provided professional development about best-grading

practices, pre-service training for teachers on best-grading practices, the belief that

teachers should be accountable for following best-grading practices, teacher-to-teacher

grading practice comparisons, district-provided common planning time for teachers in the

same content area, district-provided collaboration time for teachers on early release days,

district-provided collaboration time for teachers on late start days, district-provided

collaboration time for teachers on professional development days, teacher collaboration,

the belief that grades should be a direct reflection of student mastery of course learning

standards, grades as a means to provide feedback to stakeholders, gradebook field that

indicates student mastery of learning competencies, re-do assignments, test re-takes, and

summative and formative assessments.

RQ1: What types of support work best for teachers trying to implement best-grading practices?

RQ2: Which academic factors do teachers use when grading?

RQ3: Why do teachers choose to use academic factors when grading?

RQ4: Which non-academic factors do teachers use when grading?

RQ5: Why do teachers choose to use non-academic factors when grading?

Once survey scores were received, the researcher compared each prompts' scores using an analysis of variance (ANOVA) and used this same method for comparing teachers' demographic scores.

The quantitative data in the study investigated several areas for differences using the ANOVA test. First, a difference in the types and frequency of collaborations and interactions utilized by participants in the study was investigated through Likert-scaled survey data from questions on the survey using the ANOVA test. The researcher cross-referenced these practices to the level of implementation of best-grading practices as it pertained to the amount of time a district provided for teacher collaboration, which included professional development, conversation with administration, and collaboration with peers. The ANOVA allowed the researcher to look at the different groups and seek potential differences in perspectives.

Second, a check for differences in agreement on the topics in the survey within the participant group was investigated using the ANOVA test. Differences in agreement on the topics in the Likert-scaled survey within the participant group was investigated through survey data, topic-to-topic, using the ANOVA test. The ANOVA allowed the researcher to look at the differences in perception of agreement topic-to-topic, and

agreement-to-disagreement with survey statements, with regard to best-grading practices, measured by a Likert-scale, with regard to the support characteristics.

The researcher used a mixed methods approach for the study to investigate patterns between teacher demographic characteristics, grading, and the amount of time and support systems provided by a school district. Specific demographics that were explored included: gender, age, content area, years of service, and educational level.

When completing the quantitative analysis, several tests were used including testing for significance with the Spearman Rank Correlation Coefficient, testing the difference between two means with independent samples, testing the difference between two means with dependent samples, and running a one-sample t-test difference in means. For each test, the researcher used an alpha of 0.05 to determine the statistical significance of the results.

Limitations

While the researcher had made efforts to minimize limitations in the design of this study, such as by attempting to include a range of data points for each variable, four limitations remained. The first limitation was that the study was limited to junior/high schools in one Midwestern state. Due to this limited population, the results may not be generalizable nationwide. The second limitation was that the study included dependent variables that were self-reported data contingent upon the personal perceptions of each teacher. If teachers viewed that their perceptions of grading practices were deemed as positive or negative in the context of the study, they may have altered their responses. The third limitation was that the survey used in this study was designed to measure the perceptions of grade 7 through 12 core content area teachers' who teach in grades 7

through 12. Therefore, the results may not be generalizable to grades K through 6. The fourth limitation was that superintendents were asked to suggest the names of teachers to interview. The superintendents may have acted with bias by selecting teachers whom they felt would have positively reflected their beliefs about the appropriate relationship between grading practices and the independent variables studied.

Participants

Participants in this study were core content area teachers who taught in grades 7 through 12 at one of the four districts included in the study. For the purposes of this study, six teachers were interviewed, and all six were female teachers. According to Bluman (2008), a sample size for research must be at least 30 in order to achieve a 95% confidence interval. Thirty-four teachers completed the online survey, thereby having satisfied this criterion.

Participants who completed the survey were 41% male teachers and 59% female teachers. Teachers in this study ranged from zero to more than 15 years of teaching experience, with 50% of participants possessing 15 or more years of teaching experience. Approximately 18% of teachers possessed a bachelor's degree as their highest degree earned, 76% possessed a Master's degree, 6% possessed a specialist's degree and 0% possessed a doctorate degree. Thirty-five percent of teachers taught English/language arts, 26% taught social studies, 21% taught science, and 18% taught math. Six percent of teachers were 20 to 29 years old, 35% were 30 to 39 years old, 44% were 40 to 49 years old and 15% were 50 to 99 years old.

There were two inclusionary criteria for participation in this study: (a) the participant was a core content area teacher who taught in grades 7 through 12; and (b) the

participant was a contracted employee at one of the four Missouri SE RPDC member school districts of Washington, Adams, Jefferson, or Madison. There were two exclusionary criteria for participation in this study: (a) the teachers who were not core content area teachers who taught in grades 7 through12; and (b) teachers who were not contracted employees at one of the four Missouri SE RPDC member school districts of Washington, Adams, Jefferson, or Madison.

Conclusion

Chapter Three discussed the research design, participant descriptions, selection process, setting, statistical analysis and limitations. Chapter Four represents a discussion of the analysis of the data. Chapter Five will consists of a summary of the study, findings, conclusions, implications, and recommendations for future research.

Chapter Four: Results

This mixed methods study of grade 7th through 12th core content area teachers at four Missouri school districts focused on obtaining the demographic characteristics and insights of teachers, as well as information about their then-current grading practices, to contribute to the reform process. The research questions and null hypotheses for the study were:

RQ1: What types of support work best for teachers trying to implement best-grading practices?

RQ2: Which academic factors do teachers use when grading?

RQ3: Why do teachers choose to use academic factors when grading?

RQ4: Which non-academic factors do teachers use when grading?

RQ5: Why do teachers choose to use non-academic factors when grading?

Null H1: There will be no differences in demographic characteristics within respondents to a Likert-scale survey, with regard to best-grading practices, among the following teacher-related variables: Teachers' age, teachers' gender, teachers' levels of education, teacher's years of service, and teachers' content areas.

Null H2: There will be no differences in perception of best-grading practices, measured by a Likert-scale survey, among the following support structures: Frequency of teachers' discussions with peers regarding best-grading practices, frequency of teachers' discussions with building administrators regarding best practices, and frequency of the teachers' engagement in self-reflections in which the teacher is engaged.

Null H3: There will be no differences in perception of agreement topic-to-topic, and agreement-to-disagreement with survey statements, with regard to best-grading

practices, measured by a Likert-scale, with regard to the following support

characteristics: A student's behavior, grades as a means of punishment, student

responsibility, a student's ability level, assignment completion points, extra credit, a

teacher's personal judgment, assignments other than assessments, a student's

involvement in extracurricular activities, zeroes for missing assignments, a student's

GPA, class participation, a student's gender, a student's physical appearance, homework,

note-taking, aesthetic appearance of a student's work, a student's personality, student

effort, using grades as a reward, student progress towards achieving the course learning

standards, individual grades for group work, individual projects, project-based grading,

project-based assignments, student portfolios, conferences with students about individual

grades throughout each grading period, student-set personal learning goals for the course,

student personal reflection on achievement of the course learning standards, district-

provided time for teachers in the same content area to meet to align grading practices,

district-provided professional development about best-grading practices, pre-service

training for teachers on best-grading practices, the belief that teachers should be

accountable for following best-grading practices, teacher-to-teacher grading practice

comparisons, district-provided common planning time for teachers in the same content

area, district-provided collaboration time for teachers on early release days, district-

provided collaboration time for teachers on late start days, district-provided collaboration

time for teachers on professional development days, teacher collaboration, the belief that

grades should be a direct reflection of student mastery of course learning standards,

grades as a means to provide feedback to stakeholders, gradebook field that indicates

student mastery of learning competencies, re-do assignments, test re-takes, and

summative and formative assessments.

Null H4: There will be no relationships among levels of agreement topic-to-topic, with regard to best-grading practices, measured by a Likert-scale survey, with regard to the following support characteristics: A student's behavior, grades as a means of punishment, student responsibility, a student's ability level, assignment completion points, extra credit, a teacher's personal judgment, assignments other than assessments, a student's involvement in extracurricular activities, zeroes for missing assignments, a student's GPA, class participation, a student's gender, a student's physical appearance, homework, note-taking, the aesthetic appearance of a student's work, a student's personality, student effort, using grades as a reward, student progress towards achieving the course learning standards, individual grades for group work, individual projects, project-based grading, project-based assignments, student portfolios, conferences with students about individual grades throughout each grading period, student-set personal learning goals for the course, student personal reflection on achievement of the course learning standards, district-provided time for teachers in the same content area to meet to align grading practices, district-provided professional development about best-grading practices, pre-service training for teachers on best-grading practices, the belief that teachers should be accountable for following best-grading practices, teacher-to-teacher grading practice comparisons, district-provided common planning time for teachers in the same content area, district-provided collaboration time for teachers on early release days, district-provided collaboration time for teachers on late start days, district-provided collaboration time for teachers on professional development days, teacher collaboration, the belief that grades should be a direct reflection of student mastery of course learning

standards, grades as a means to provide feedback to stakeholders, gradebook field that

indicates student mastery of learning competencies, re-do assignments, test re-takes, and

summative and formative assessments.

General Qualitative Feedback

As previously discussed, the purpose of this study was to investigate patterns

between teacher demographic characteristics, grading, and the amount of time and

support systems provided by a school district. Specific demographics explored included:

gender, age, content area, years of service, and educational level. Chapter Five explains

the findings, or results, of six qualitative interviews and 34 survey responses in narrative

form. In Chapter Four, an analysis of the data is presented in a descriptive format,

including tables.

Demographics

To check for demographic characteristics within the participant group, the

researcher considered Null Hypothesis 1: There are no differences in demographic

characteristics within respondents to a Likert-scale survey, with regard to best-grading

practices, among the following teacher-related variables: Teachers' age, teachers' gender,

teachers' levels of education, teachers' years of service, and teachers' content areas. With

regard to demographics of study participants, a z-test for difference in proportions was

applied to check for differences in outstanding demographic characteristics that could

possibly influence the responses to survey questions.

Age

Table 1 shows the data for the number of participants representing categories of

age among the teacher participants. For this characteristic, the null sub-hypothesis, there

is no difference in age representation among participants, was rejected for two

comparison pairings and not rejected for a third pairing. There were significant

differences between the number of 30-somethings compared to the number of 20-

somethings (z = -2.9990; z-critical = 1.96), and between the number of 40-somethings

compared to the number of 30- somethings (z = -2.6100; z-critical = 1.96). However,

there was no significant difference between the number of 40- somethings compared to

the number of 50-somethings (z = -0.7440; z-critical = 1.96).

Table 1

Prompt 5- Age

#	Answer	%	Count
1	20-29 years old	5.88%	2
2	30-39 years old	35.29%	12
3	40-49 years old	44.12%	15
4	50-99 years old	14.71%	5
	Total	100%	34

Gender

Table 2 shows the data for the number of participants representing male and

female. For this characteristic, the null sub-hypothesis, there is no difference in gender

representation, was not rejected (z = -1.455; z-critical = 1.96). Though observably more

female teachers were employed, the findings were not significantly different and the

percentage of one gender was not significantly more than the other gender.

Table 2

Prompt 2 – Gender

#	Answer	%	Count
1	Male	41.18%	14
2	Female	58.82%	20
	Total	100%	34

Level of Education

Table 3 shows the data for the percentage of teachers holding each of the various educational degree levels among the participants. For this characteristic, the null sub-hypothesis, there is no difference in the percentage of teachers earning degrees higher than a bachelor's, was rejected (z = -4.859; z-critical = 1.96). The number of teachers with a master's degree was observably higher, and also significantly higher, than the number of teachers with a bachelor's degree.

Table 3

Prompt 3 – Highest Degree Earned

#	Answer	%	Count
1	Bachelor's	17.65%	6
2	Master's	76.47%	26
3	Specialist	5.88%	2
4	Doctorate	0.00%	0
	Total	100%	34

Years of Service

Table 4 shows the data for the number of years of teaching experience among the participants. For this characteristic, the null sub-hypothesis, there is no difference in the number of years of teaching experience, was not rejected (z = -1.736; z-critical = ±1.96). Though the number of years of experience was observably more frequent on the higher end of the range, the differences in years of experience throughout were not significant.

Table 4

Prompt 1 - Years of Teaching Experience

#	Answer	%	Count
1	0-5	11.76%	4
2	6-10	8.82%	3
3	11-15	29.41%	10
4	15+	50.00%	17
	Total	100%	34

Teaching Content Area

Table 5 shows the data for the number of participants representing various content areas in the educational setting. For this characteristic, the null sub-hypothesis, there is no difference in representation among content areas, was not rejected ($z = -0.787$; z-critical = 1.96). There were no significant differences in the categories of content area represented among the participants; therefore, no one content area was influencing the outcomes of the perceptions measured for this study.

Table 5

Prompt 4 - Content Area That You Teach

#	Answer	%	Count
1	English/Language Arts	35.29%	12
2	Social Studies	26.47%	9
3	Science	20.59%	7
4	Math	17.65%	6
	Total	100%	34

Types of Interactions

Null Hypothesis 2 considered the types and frequency of collaborations and interactions utilized by participants in the study. To check for the types and frequency of collaborations by the participant group, the researcher considered null sub-hypothesis 2: There are no differences in perceptions of best-grading practices, measured by a Likert-scale survey, among the following support structures: Frequency of teacher's discussions with peers regarding best-grading practices, frequency of teacher's discussions with building administrators regarding best practices, and frequency of the teacher's engagement in self-reflection in which the teacher is engaged.

To check for significant differences in the frequency of three different types of collaboration, an ANOVA was applied to data for survey prompts 43, 58, and 60, which

were 'How often does your administrator engage in dialogue with you about best-grading practices?,' 'How often do you collaborate with your peers?,' and 'How often do you self-reflect about the accuracy and fairness of your grading practices?' The data are displayed on Table 6.

Table 6

Frequency of the Use of Each Type of Interaction (Prompts 42, 58, 60)

	admin dialogue	peer collaboration	self-reflection
Never	11.76%	8.82%	8.82%
1 x daily	0.00%	23.53%	26.47%
1 x weekly	5.88%	23.53%	35.29%
1 x monthly	14.71%	23.53%	2.94%
1 x quarterly	26.47%	20.59%	14.71%
1 x per semester	26.47%	0.00%	5.88%
1 x yearly	14.71%	0.00%	5.88%

Results of the ANOVA, when checking the frequency of the type of collaboration, are displayed on Table 7.

Table 7

ANOVA: Types of Collaboration with Peers (Prompt 58)

Groups	Count	Sum	Average	Variance
admin dialogue	7	1	0.1429	0.0096
peer collaboration	7	1	0.1429	0.0122
self-reflection	7	1	0.1428	0.0148

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	9.52E-10	2	4.76E-10	3.89E-08	1	3.5546
Within Groups	0.2202	18	0.0122			
Total	0.2202	20				

In considering administrative dialogue, peer collaboration, and self-reflection as types of collaboration, with regard to the null sub-hypothesis, there are no differences in

the types of collaboration utilized; the researcher did not reject the null hypothesis and did not support the hypothesis ($\alpha = .05$; $p = 1$; $F = .0000$; F-critical $= 3.5546$). There were no significant differences in the types of collaboration in which teachers participated.

Results of the ANOVA when checking the frequency of the use of self-reflection in grading practices are displayed on Table 8.

Table 8

ANOVA: Frequency of Self-Reflection When Considering Grading Practices (Prompt 60)

Groups	Count	Sum	Average	Variance
Never	3	0.2940	0.0980	0.0003
1 x daily	3	0.5000	0.1667	0.0210
1 x weekly	3	0.6470	0.2157	0.0219
1 x monthly	3	0.4118	0.1373	0.0107
1 x quarterly	3	0.6177	0.2059	0.0035
1 x per semester	3	0.3235	0.1078	0.0193
1 x yearly	3	0.2059	0.0686	0.0055

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	0.0559	6	0.0093	0.7931	0.5904	2.8477
Within Groups	0.1644	14	0.0117			
Total	0.2202	20				

In considering the frequency of self-reflection in use of grading practices, with regard to the null sub-hypothesis, there are no differences in the frequency of self-reflection, the researcher did not reject the null hypothesis and did not support the hypothesis ($\alpha = .05$; $p = .5904$; $F = .7931$; F-critical $= 2.8477$). There were no significant differences in the frequency with which teachers applied self-reflection to their own grading practices.

A Check for Differences

To check for differences in agreement on the topics in the survey within the

participant group, the researcher considered Null Hypothesis 3: There are no differences in perception of agreement topic-to-topic, and agreement-to-disagreement with survey statements, with regard to best-grading practices, measured by a Likert-scale, with regard to the following support characteristics: A student's behavior, grades as a means of punishment, student responsibility, a student's ability level, assignment completion points, extra credit, a teacher's personal judgment, assignments other than assessments, a student's involvement in extracurricular activities, zeroes for missing assignments, a student's GPA class participation, a student's gender, a student's physical appearance, homework, note-taking, aesthetic appearance of a student's work, a student's personality, student effort, using grades as a reward, student progress towards achieving the course learning standards, individual grades for group work, individual projects, project-based grading, project-based assignments, student portfolios, conferences with students about individual grades throughout each grading period, student-set personal learning goals for the course, student personal reflection on achievement of the course learning standards, district-provided time for teachers in the same content area to meet to align grading practices, district-provided professional development about best-grading practices, pre-service training for teachers on best-grading practices, the belief that teachers should be accountable for following best-grading practices, teacher-to-teacher grading practice comparisons, district-provided common planning time for teachers in the same content area, district-provided collaboration time for teachers on early release days, district-provided collaboration time for teachers on late start days, district-provided collaboration time for teachers on professional development days, teacher collaboration, the belief that grades should be a direct reflection of student mastery of course learning standards,

grades as a means to provide feedback to stakeholders, gradebook field that indicates

student mastery of learning competencies, re-do assignments, test re-takes, and

summative and formative assessments.

An ANOVA was applied to check for differences in agreement on the topics

listed in Null Hypothesis 3. No differences were established; therefore, no individual *t*-

tests or *z*-tests were applied. The analysis discussion follows.

Topic-to-Topic

Concerning the analysis for agreement versus disagreement, topic-to-topic, on

survey prompts 6 through 56, no significant differences were found when ANOVA was

applied ($\alpha = .05$; $p = 1.0000$; $F = .0000$; F-critical $= 1.1480$). Note, Table 1A (see

Appendix L) displays ANOVA results. With regard to the null sub-hypothesis, there are

no differences in perceptions, with regard to survey prompts, when considering topic-to-

topic discussion; the researcher did not reject the null hypothesis. No single topic or set of

topics stood out as having stronger perceptions than the entire grouping of survey

prompts considered. Therefore, no individual *z*-tests for difference in proportion were

applied, with regard to topic-to-topic discussion.

Agreement Versus Disagreement, Question-to-Question

Analysis of responses to survey prompts 6 through 41 and 43 through 56 was

completed by applying a *z*-test for difference in proportions to the percentage of

agreement and percentage of disagreement to each prompt. Agreement was represented

by a combination of agree and strongly agree responses. Disagreement was represented

by a combination of disagree and strongly disagree responses.

ANOVA was applied to the percentage of agreement versus the percentage of disagreement for survey prompts 6 through 41, question-to-question, to identify whether strong differences in perception existed. Significant differences were not found when ANOVA was applied ($\alpha = .05$; $p = 1.0000$; $F = .0000$; F-critical $= 1.5050$). With regard to the null sub-hypothesis, there are no differences in perceptions, with regard to survey prompts, when considering topic-to-topic discussion; the researcher did not reject the null hypothesis. No significant differences were established, question-to-question.

Concerning the analysis for agreement versus disagreement topic-to-topic on survey prompts 43 through 56, no significant differences were found when ANOVA was applied (hypothesis ($\alpha = .05$; $p = 1.0000$; $F = .0000$; F-critical $= 1.8992$). With regard to the null sub-hypothesis, there are no differences in perceptions with regard to survey prompts, when considering topic-to-topic discussion, the researcher did not reject the null hypothesis. No significant differences were established, question-to-question. No single topic or set of topics stood out as having stronger perceptions than the entire grouping of survey prompts considered. Therefore, no individual z-tests for difference in proportion were applied, with regard to topic-to-topic, or question-to-question discussion.

Agreement Versus Disagreement, Within Each Question

To apply Null Hypothesis 3 to perception topics represented by individual questions, a z-test for difference in proportions was applied to check for differences in agreement versus disagreement on the topics covered by individual questions and listed in Null Hypothesis 3. For each individual question, the strongly agree and agree percentages were added together and the strongly disagree and disagree percentages were added together. This left the percentage of participants who answered '3' to the question

prompt out of the analysis, allowing for a neutral column, concerning perception.

Table 9 displays results for question prompts 6 through 41. The following question prompts indicated significant disagreement with the prompt: 6, 7, 8, 14, 17, 20, 24, and 37. Complete disagreement was indicated by participants for prompts 19, 21, and 26.

Significant agreement was indicated by participants for the following prompts: 9, 11, 12, 13, 16, 18, 22, 25, 27, 28, 29, 30, 31, 32, 33, 35, 36, and 41 (z-test values are included on Table 9). There were no significant differences found for prompts 10, 15, 23, 34, 38, 39, and 40.

Table 9

Individual Question Agreement versus Disagreement (Prompt 6-41)

	#	z-test value		#	z-test value
significant disagreement	6	-3.908	significant disagreement	24	-5.114
significant disagreement	7	-7.288	significant agreement	25	2.429
significant disagreement	8	-5.114	complete disagreement	26	
significant agreement	9	2.190	significant agreement	27	4.397
no significance	10	1.497	significant agreement	28	4.885
significant agreement	11	3.908	significant agreement	29	5.640
significant agreement	12	4.397	significant agreement	30	5.114
significant agreement	13	5.581	significant agreement	31	6.575
significant disagreement	14	-6.803	significant agreement	32	3.420
no significance	15	1.479	significant agreement	33	4.885
significant agreement	16	7.522	no significance	34	-1.555
significant disagreement	17	-8.007	significant agreement	35	5.600
significant agreement	18	4.659	significant agreement	36	2.593
complete disagreement	19		significant disagreement	37	-2.538
significant disagreement	20	-5.863	no significance	38	1.240
complete disagreement	21		no significance	39	0.493
significant agreement	22	5.582	no significance	40	0.256
no significance	23	1.767	significant agreement	41	5.282

Note. Critical value = 1.96.

Table 10 displays results for question prompts 43 through 56. The following

question prompts indicated significant disagreement with the prompt: 44, 45, 46, and 51.

Significant agreement was indicated by participants for prompts 43, 47, 49, 50, 52, 53, 55, and 56 (z-test values are included on the Table 10). There were no significant differences found for prompts 48 and 54.

Table 10

Individual Question Agreement versus Disagreement (Prompt 43-56)

significant agreement	43	2.995
significant disagreement	44	-1.945
significant disagreement	45	-1.971
significant disagreement	46	-1.993
significant agreement	47	5.831
no significance	48	0
significant agreement	49	7.061
significant agreement	50	5.989
significant disagreement	51	-2.264
significant agreement	52	4.000
significant agreement	53	4.138
no significance	54	-1.523
significant agreement	55	4.148
significant agreement	56	3.111

Note. Critical value = 1.96.

Significant disagreement was found in participant responses to the following prompts:

6: I consider a student's behavior when grading (z-test value = -3.908; z-critical value = ±1.96).

7: Grades are a means to punish students for low academic achievement (z-test value = -7.288; z-critical value = ±1.96).

8: I assign points for bringing daily supplies to class (z-test value = -5.114; z-critical value = ±1.96).

14: I consider a student's involvement in extracurricular activities when grading (z-test value = -6.803; z-critical value = ±1.96).

17: I consider a student's GPA when grading (z-test value = -8.007; z-critical value = ±1.96).

20: I have lowered a student's grade for demonstrating poor behavior (z-test value = -5.863; z-critical value = ±1.96).

24: I consider the aesthetic appearance of the students' work when grading (z-test value = -5.114; z-critical value = ±1.96).

37: I require my students to reflect on their achievement of the course learning standards in a journal or notebook (z-test value = -2.538; z-critical value = ±1.96).

As seen in the z-test values listed on Table 10, significant agreement was found in participant responses to the following prompts:

9: I consider a student's ability level when grading (z-test value = 2.190; z-critical value = ±1.96).

11: I allow students to earn extra credit (z-test value = 3.908; z-critical value = ±1.96).

12: Teachers should be able to exercise a degree of personal judgment in assigning grades (z-test value = 4.397; z-critical value = ±1.96).

13: I assign grades for assignments other than assessments (z-test value = 5.581; z-critical value = ±1.96).

16: Students lose points if they do not complete all parts of an assignment (z-test value = 7.522; z-critical value = ±1.96).

18: I assign grades for class participation (z-test value = 4.659; z-critical value = ±1.96).

22: I assign grades for homework (z-test value = 5.582; z-critical value = ±1.96).

25: I assign students a lower grade for turning in an assignment after its due date (z-test value = 2.429; z-critical value = ±1.96).

27: I consider a student's effort when grading (z-test value = 4.397; z-critical value = ±1.96).

28: Grades should reward students for high academic achievement (z-test value = 4.885; z-critical value = ±1.96).

29: I consider student progress towards achieving the course learning standards when grading (z-test value = 5.640; z-critical value = ±1.96).

30: I assign individual grades for students working in groups (z-test value = 5.114; z-critical value = ±1.96).

31: I assign grades for individual projects (z-test value = 6.575; z-critical value = ±1.96).

32: I use project-based grading in my classroom (z-test value = 3.420; z-critical value = ±1.96).

33: I assign grades for project-based assignments (z-test value = 4.885; z-critical value = ±1.96).

35: I conference with students about their grades throughout each grading period (z-test value = 5.600; z-critical value = ±1.96).

36: I require students to set personal learning goals for the course (z-test value = 2.593; z-critical value = ±1.96).

41: Teachers should be accountable for following best-grading practices (z-test value = 5.282; z-critical value = ±1.96).

As seen in the z-test values listed on Table 9 and Table 10, no significant agreement or disagreement was found in participant responses to prompts

10: I assign points for completing an assignment (z-test value = 1.497; z-critical value = ±1.96).

15: I assign zeroes for missing assignments (z-test value = 1.479; z-critical value = ±1.96).

23: I assign grades for taking notes (z-test value = 1.767; z-critical value = ±1.96).

34: I include student portfolios in my grading (z-test value = -1.555; z-critical value = ±1.96).

38: My district provides time for teachers in the same content area to meet to align grading practices (z-test value = 1.240; z-critical value = ±1.96).

39: My district provides professional development about best-grading practices (z-test value = 0.493; z-critical value = ±1.96).

40: I received training on best-grading practices in my undergraduate coursework (z-test value = 0.256; z-critical value = ±1.96).

In addition, complete disagreement was found with prompts 19, 21, and 26.

19: I consider a student's gender when grading.

21: I consider the physical appearance of a student when grading.

26: I consider a student's personality when grading.

Though trends of strong agreement were found in responses that allowed a grouping of prompts that received similar response patterns from participants, there were

no significant differences found in a comparison of the level of response rates when
considered topic-to-topic (p value = 0.9999; alpha value = 0.05; F-test value =
0.000000000175; F-critical value = 1.4180).

Strong Positive Relationships

For prompts 6 through 41 and 43 through 56, potential relationships between
participant responses were analyzed with the application of the Pearson Product Moment
Correlation Coefficient (PPMCC). Analysis on responses to the prompts represented by
strongly disagree, disagree, neither agree nor disagree, somewhat agree, and strongly
agree yielded many significant relationships (see Appendix M). This report discusses the
relationships found to be strongly and positively related, with an r value greater than
0.750.

Strong relationships were found between a number of prompts in the study
survey. Prompts are grouped together according to strong agreement in response rates
among the topics presented in the survey (see Table 11). For example, participants
responded with significant disagreement to prompt 6. Responses to prompts 7, 8, 14, 17,
19, 20, 21, 24, and 26 indicated the same trends of agreement as prompt 6; when strongly
disagree was the strongest percentage in 6, strongly disagree was also the strongest
percentage in 7, 8, 14, 17, 19, 20, 21, 24, and 26.

Table 11

Strong Positive Relationships in Survey Responses

Q	Strong Positive Relationship (> 0.750) With
# 6	7, 8, 14, 17, 29, 02, 21, 24, 26
# 7	6, 8, 14, 17, 19, 20, 21, 24, 26
# 8	14, 17, 19, 20, 21, 24, 26
# 9	10, 11, 12, 30, 32
# 10	11, 12, 23, 30, 36
# 11	9, 10, 12, 18, 23, 25, 29, 30, 32, 39
# 12	9, 10, 11, 18, 25, 27, 29, 30, 32, 39
# 13	16, 22, 27, 28, 29, 31, 33, 35, 41
# 14	6, 7, 8, 19, 20, 21, 24, 26
# 15	35
# 16	13, 22, 28, 31, 33, 35, 41
# 17	6, 7, 8, 14, 19, 20, 21, 24, 26
# 18	11, 12, 22, 25, 27, 28, 29, 30, 32, 33, 41
# 19	6, 7, 8, 14, 17, 20, 21, 24, 26
# 20	6, 7, 8, 14, 17, 19, 21, 24, 26
# 21	6, 7, 8, 14, 17, 19, 20, 24, 26
# 22	13, 16, 18, 25, 27, 28, 29, 30, 32, 33, 35
# 23	10, 11, 36, 39
# 24	6, 7, 8, 14, 17, 19, 20, 21, 26
# 25	11, 12, 18, 22, 27, 29, 30, 32
# 26	6, 7, 8, 14, 17, 19, 20, 21
# 27	12, 13, 18, 22, 25, 28, 29, 30, 32, 33, 35, 41
# 28	14, 16, 18, 22, 27, 29, 31, 33, 35, 41
# 29	11, 12, 13, 18, 22, 25, 28, 30, 32, 33, 35
# 30	9, 10, 11, 12, 18, 22, 25, 27, 29, 32, 39
# 31	13, 16, 28, 33, 35, 41
# 32	9, 11, 12, 18, 22, 25, 28, 29, 31, 33
# 33	13, 16, 18, 22, 27, 28, 29, 32, 35, 41
# 34	31
# 35	13, 15, 16, 27, 28, 29, 31, 33, 41
# 36	10, 23
# 37	34
# 38	
# 39	11, 12, 23, 30

Table 11. Continued.

# 40	
# 41	13, 16, 18, 27, 28, 29, 31, 33, 35
# 42	
# 43	49, 55
# 44	45, 48
# 45	44, 46, 51
# 46	45, 51
# 47	50, 52
# 48	44, 51
# 49	43, 50, 52
# 50	47, 49, 52
# 51	45, 46, 48
# 52	47, 49, 50
# 53	56
# 54	
# 55	43
# 56	53

Table 12 lists the prompt numbers next to each question asked on the survey.

Please refer to the table for details pertaining to the synthesis of each prompt.

Table 12

Survey Prompts

1. Years of teaching experience:	0-5	5-10	10-15	15+
2. Gender:	M	F		
3. Highest degree earned:	BS	Master	Specialist	Doctorate
4. Content area:	ELA	S.S.	Science	Math
5. Age:	20-29	30-39	40-49	50+
6. I consider a student's behavior when grading				
7. Grades are a means to punish students for low academic achievement				
8. I assign points for bringing daily supplies to class				
9. I consider a student's ability level when grading				
10. I assign points just for completing an assignment				
11. I allow students to earn extra credit				
12. Teachers should be able to exercise a degree of personal judgment in assigning grades				
13. I assign grades for assignments other than assessments				
14. I consider a student's involvement in extracurricular activities when grading				
15. I assign students zeroes for missing assignments				

Continued

Table 12 Continued

16. Students lose points if they do not complete all parts of an assignment
17. I consider a student's GPA when grading
18. I assign grades for class participation
19. I consider a student's gender when grading
20. I have lowered a student's grade for demonstrating poor behavior
21. I consider the physical appearance of a student when grading
22. I assign grades for homework
23. I assign grades for taking notes
24. I consider the aesthetic appearance of the students' work when grading
25. I assign students a lower grade for turning in an assignment after its due date
26. I consider a student's personality when grading
27. I consider a student's effort when grading
28. Grades should reward students for high academic achievement
29. I consider student progress towards achieving the course learning standards when grading
30. I assign individual grades for students working in groups
31. I assign grades for individual projects
32. I use project-based grading in my classroom
33. I assign grades for project-based assignments
34. I include student portfolios in my grading
35. I conference with students about their grades throughout each grading period
36. I require students to set personal learning goals for the course
37. I require my students to reflect on their achievement of the course learning standards in a journal or notebook
38. My district provides time for teachers in the same content area to meet to align grading practices
39. My district provides professional development about best-grading practices
40. I received training on best-grading practices in my undergraduate coursework
41. Teachers should be accountable for following best-grading practices
42. How often does your administrator engage in dialogue with you about best-grading practices? Never, 1x daily, 1x weekly, 1x monthly, 1x quarterly, 1x per semester, or 1x yearly
43. I consider how my grading practices compare to other teachers' grading practices in the same content area as myself
44. My district provides common planning time for teachers in the same content area to promote collaboration
45. My district provides collaboration time for teachers on early release days
46. My district provides collaboration time for teachers on late start days
47. My district provides collaboration time for teachers on professional development days
48. I regularly collaborate with grade level teachers during common plan time
49. I believe grades should be a direct reflection of student mastery of learning standards
50. My grades provide feedback about student progress of the learning standards to stakeholders

Table 12 Continued

51. My gradebook includes a field that indicates each student's mastery of learning competencies
52. I allow students to re-do assignments
53. I allow students to re-take a test if they fail it
54. I grade summative assessments, not formative assessments
55. I believe that my grades are an indication of how students should perform on state exams
56. I weigh summative assessments heavier than formative assessments when grading
57. What other factors do you consider when grading a student's work?
58. How often do you collaborate with your peers? (Please select one response.) Never, 1x daily, 1x weekly, 1x monthly, 1x quarterly, 1x per semester, or 1x yearly
59. What topics do you discuss with your peers if you collaborate with them?
60. How often do you self-reflect about the accuracy and fairness of your grading practices? (Please select one response.) Never, 1x daily, 1x weekly, 1x monthly, 1x quarterly, 1x per semester, or 1x yearly

Synthesis of the strong positive relationship with prompt 6 included the following non-academic prompts: 7, 8, 14, 17, 19, 20, 21, 24, and 26. After analyzing the relationship with prompt 6, it was shown that participants had a strong disagreement in considering the following non-academic factors when grading: using grades as a means to punish students for low academic achievement, bringing daily supplies to class, a student's involvement in extracurricular activities, a student's GPA, a student's physical appearance, the aesthetic appearance of a student's work, and a student's personality.

Synthesis of the strong positive relationship with prompt 7 included the following non-academic prompts: 6, 8, 14, 17, 19, 20, 21, 24, and 26. After analyzing the relationship with prompt 7, it was shown that participants had a strong disagreement in considering the following non-academic factors when grading: a student's behavior, using grades as a means to punish students for low academic achievement, bringing daily supplies to class, a student's involvement in extracurricular activities, a student's GPA, a student's gender, a student's physical appearance, the aesthetic appearance of a student's work, and a student's personality.

Synthesis of the strong positive relationship with prompt 8 included the following non-academic prompts: 14, 17, 19, 20, 21, 24, and 26. After analyzing the relationship with prompt 8, it was shown that participants had a strong disagreement in considering the following non-academic factors when grading: a student's involvement in extracurricular activities, a student's GPA, a student's gender, a student's behavior, a student's physical appearance, the aesthetic appearance of a student's work, and a student's personality.

Synthesis of the strong positive relationship with prompt 9 included the following non-academic prompts: 10, 11, and 12 and academic prompts 30 and 32. After analyzing the relationship with prompt 9, it was shown that participants included both academic and non-academic factors in grades. Participants showed a strong agreement with the inclusion of the following non-academic factors when grading: completion points, extra credit, and teachers' ability to exercise a degree of personal judgment in assigning grades. Participants also showed a strong agreement with the inclusion of two academic factors when grading. This included the assignment of individual grades to each student when working in groups and using project-based grading in classrooms.

Synthesis of the strong positive relationship with prompt 10 included the following non-academic prompts: 11, 12, and 23 and academic prompts 30 and 36. After analyzing the relationship with prompt 10, it was shown that participants included both academic and non-academic factors in grades. Participants showed a strong agreement with the inclusion of the following non-academic factors when grading: Extra credit, teachers' ability to exercise a degree of personal judgment in assigning grades, and note-taking. Participants also showed a strong agreement with the inclusion of the following

academic factors when grading: the assignment of individual grades to each student when working in groups and requiring student to set personal learning goals for the course.

Synthesis of the strong positive relationship with prompt 11 included the following non-academic prompts: 9, 10, 12, 18, 23, and 25 and the following academic prompts: 29, 30, 32, and 39. After analyzing the relationship with prompt 11, it was shown that participants included both academic and non-academic factors in grades. Participants showed a strong agreement with the inclusion of the following non-academic factors when grading: Students' ability levels, assignment completion points, a degree of personal judgment on behalf of teachers, class participation, note-taking, and the lowering of grades for turning an assignment in late. Participants also showed a strong agreement with the inclusion of the following academic factors when grading: a student's progress towards achieving the course learning standards, individual grades for students working in groups, the use project-based grading in their classroom, and that their districts provided professional development about best-grading practices.

Synthesis of the strong positive relationship with prompt 12 included the following non-academic prompts: 9, 10, 11, 18, 25, and 27 and the following academic prompts: 29, 30, 32, and 39. After analyzing the relationship with prompt 12, it was shown that participants included both academic and non-academic factors in grades. Participants showed a strong agreement with the inclusion of the following non-academic factors when grading: the teacher's consideration of a student's ability level, assignment completion points, allowing students to earn extra credit, class participation, lowering a student's grade for turning an assignment after its due date, and a student's effort. Participants showed a strong agreement with the inclusion of the following academic

factors when grading: a student's progress towards achieving the course learning standards, assigning individual grades for students working in groups, project-based grading, and district-provided professional development about best-grading practices.

Synthesis of the strong positive relationship with prompt 13 included the following non-academic prompts: 16, 22, 27, and 28 and the following academic prompts: 29, 31, 33, 35, and 41. After analyzing the relationship with prompt 13, it was shown that participants included both academic and non-academic factors in grades. Participants showed a strong agreement with the inclusion of the following non-academic factors when grading: a student losing points for failing to complete all parts of an assignment, homework, a student's effort, and the use of grades as a reward for a student's high academic achievement. Participants showed a strong agreement with the inclusion of the following academic factors when grading: a student's progress towards achieving the course learning standards, individual projects, project-based assignments, conferences with students about grades, and the belief that teachers should be accountable for following best-grading practices.

Synthesis of the strong positive relationship with prompt 14 included the following non-academic prompts: 6, 7, 8, 19, 20, 21, 24, and 26. After analyzing the relationship with prompt 14, it was shown that participants agreed that particular non-academic factors should not be included in grades. Participants showed a strong disagreement with the inclusion of the following non-academic factors when grading: a student's behavior, using grades as punishment for low academic achievement, bringing daily supplies to class, a student's gender, a student's physical appearance, the aesthetic appearance of a student's work, and a student's personality.

Synthesis of the strong positive relationship with prompt 15 included academic prompt 35. After analyzing the relationship with prompt 15, it was shown that participants included an academic factor in grades. Participants showed a strong agreement with the inclusion of the academic factor of conferencing with a student about their grades throughout each grading period when grading.

Synthesis of the strong positive relationship with prompt 16 included the following non-academic prompts: 13, 22, and 28 and the following academic prompts: 31, 33, 35, and 41. After analyzing the relationship with prompt 16, it was shown that participants included both academic and non-academic factors in grades. Participants showed a strong agreement with the inclusion of the following non-academic factors when grading: assignments other than summative assessments, homework, and using grades as a reward for high academic achievement. Participants showed a strong agreement with the inclusion of the following academic factors when grading: individual projects, project-based assignments, conferencing with a student about their grades throughout each grading period, and the belief that teachers should be accountable for following best-grading practices.

Synthesis of the strong positive relationship with prompt 17 included the following non-academic prompts: 6, 7, 8, 14, 19, 20, 21, 24, and 26. After analyzing the relationship with prompt 17, it was shown that participants agreed that particular non-academic factors should not be considered in grades. Participants showed a strong disagreement with the inclusion of the following non-academic factors when grading: a student's behavior, grades as a means of punishment for a student's low academic achievement, bringing daily supplies to class, a student's involvement in extracurricular

activities, a student's gender, a student's physical appearance, the aesthetic appearance of a student's work, and a student's personality.

Synthesis of the strong positive relationship with prompt 18 included the following non-academic prompts: 11, 12, 22, 25, 27, and 28 and the following academic prompts: 29, 30, 32, 33, and 41. After analyzing the relationship with prompt 18, it was shown that participants included both academic and non-academic factors in grades. Participants showed a strong agreement with the inclusion of the following non-academic factors when grading: extra credit, a teacher's personal judgment, homework, lowering a student's grade for turning in an assignment after its due date, a student's effort, and using grades as a reward for high academic achievement. Participants showed a strong agreement with the inclusion of the following academic factors when grading: a student's progress towards achieving the course learning standards, individual grades for group work, project-based grading, and the belief that teachers should be accountable for following best-grading practices.

Synthesis of the strong positive relationship with prompt 19 included the following non-academic prompts: 6, 7, 8, 14, 17, 20, 21, 24, and 26. After analyzing the relationship with prompt 19, it was shown that participants disagreed with the inclusion of particular non-academic factors in grades. Participants showed a strong disagreement with the inclusion of the following non-academic factors when grading: a student's behavior, using grades as punishment for a student's low academic achievement, bringing daily supplies to class, a student's involvement in extracurricular activities, a student's GPA, a student's physical appearance, the aesthetic appearance of a student's work, and a student's personality.

Synthesis of the strong positive relationship with prompt 20 included the following non-academic prompts: 6, 7, 8, 14, 17, 19, 21, 24, and 26. After analyzing the relationship with prompt 20, it was shown that participants disagreed with the inclusion of particular non-academic factors in grades. Participants showed a strong disagreement with the inclusion of the following non-academic factors when grading: a student's behavior, using grades as punishment for a student's low academic achievement, bringing daily supplies to class, a student's involvement in extracurricular activities, a student's GPA, a student's gender, a student's physical appearance, the aesthetic appearance of a student's work, and a student's personality.

Synthesis of the strong positive relationship with prompt 21 included the following non-academic prompts: 6, 7, 8, 14, 17, 19, 20, 24, and 26. After analyzing the relationship with prompt 21, it was shown that participants disagreed with the inclusion of particular non-academic factors in grades. Participants showed a strong disagreement with the inclusion of the following non-academic factors when grading: a student's behavior, using grades as punishment for a student's low academic achievement, bringing daily supplies to class, a student's involvement in extracurricular activities, a student's GPA, a student's gender, a student's physical appearance, the aesthetic appearance of a student's work, and a student's personality.

Synthesis of the strong positive relationship with prompt 22 included the following non-academic prompts: 13, 16, 18, 25, 27, and 28 and academic prompts: 29, 30, 32, 33, and 35. After analyzing the relationship with prompt 22, it was shown that participants included both academic and non-academic factors in grades. Participants showed a strong agreement with the inclusion of the following non-academic factors

when grading: assignments other than summative assessments, completion points, class participation, turning in an assignment after its due date, and the belief that grades should reward students for high academic achievement. Participants showed a strong agreement with the inclusion of the following academic factors: a student's progress towards achieving the course learning standards, individual grades for group work, project-based assignments, and the act of conferencing with a student about their grades throughout each grading period.

Synthesis of the strong positive relationship with prompt 23 included non-academic prompts 10 and 11 and academic prompts 36 and 39. After analyzing the relationship with prompt 23, it was shown that participants included both academic and non-academic factors in grades. Participants showed a strong agreement with the inclusion of the non-academic factors of completion points and extra credit when grading. Participants showed a strong agreement with the inclusion of the academic factors of requiring students to set personal course learning goals and district-provided professional development about best-grading practices when grading.

Synthesis of the strong positive relationship with prompt 24 included the following non-academic prompts: 6, 7, 8, 14, 17, 19, 20, 21, and 26. After analyzing the relationship with prompt 24, it was shown that participants disagreed with the inclusion of particular non-academic factors in grades. Participants showed a strong disagreement with the inclusion of the following non-academic factors when grading: a student's behavior, using grades as punishment for a student's low academic achievement, bringing daily supplies to class, a student's involvement in extracurricular activities, a student's GPA, a student's physical appearance, and a student's personality.

Synthesis of the strong positive relationship with prompt 25 included the following non-academic prompts: 11, 12, 18, 22, and 27 and the following academic prompts: 29, 30, and 32. After analyzing the relationship with prompt 25, it was shown that participants included both academic and non-academic factors in grades. Participants showed a strong agreement with the inclusion of the following non-academic factors when grading: extra credit, a teacher's personal judgment, class participation, homework, and a student's effort. Participants showed a strong agreement with the inclusion of the following academic factors when grading: a student's progress towards achieving the course learning standards, individual grades for group work, and project-based assignments.

Synthesis of the strong positive relationship with 26 included the following non-academic prompts: 6, 7, 8, 14, 17, 19, 20, and 21. After analyzing the relationship with prompt 26, it was shown that participants disagreed about the inclusion of particular non-academic factors in grades. Participants showed a strong disagreement with the inclusion of the following non-academic factors when grading: a student's behavior, using grades as punishment for a student's low academic achievement, bringing daily supplies to class, a student's involvement in extracurricular activities, a student's GPA, a student's gender, and a student's physical appearance.

Synthesis of the strong positive relationship with 27 included the following non-academic prompts: 12, 13, 18, 22, 25, and 28 and the following academic prompts: 29, 30, 32, 33, 35, and 41. After analyzing the relationship with prompt 27, it was shown that participants included both academic and non-academic factors in grades. Participants showed a strong agreement with the inclusion of the following non-academic factors

when grading: a teacher's personal judgment, assignments other than summative

assessments, class participation, homework, turning in an assignment after its due date,

and using grades to reward a student's high academic achievement. Participants showed a

strong agreement with the inclusion of the following academic factors when grading: a

student's progress towards achieving the course learning standards, individual grades for

group work, project-based assignments, the act of conferencing with a student about their

grades throughout each grading period, and the belief that teachers should be accountable

for following best-grading practices.

Synthesis of the strong positive relationship with prompt 28 included the

following non-academic prompts: 14, 16, 18, 22, and 27 and the following academic

prompts: 29, 31, 33, 35, and 41. After analyzing the relationship with prompt 28, it was

shown that participants included both academic and non-academic factors in grades.

Participants showed a strong agreement with the inclusion of the following non-academic

factors when grading: point deduction for failing to complete all parts of an assignment,

class participation points, homework, and a student's effort. However, participants

showed a strong disagreement with the inclusion of a student's involvement in

extracurricular activities. Participants showed a strong agreement with the inclusion of

the following academic factors when grading: a student's progress towards achieving the

course learning standards, individual projects, project-based assessments, the act of

holding student conferences with students about their grades throughout each grading

period, and the belief that teachers should be accountable for following best-grading

practices.

Synthesis of the strong positive relationship with prompt 29 included the following non-academic prompts: 11, 12, 13, 18, 22, 25, and 28 and the following academic prompts: 30, 32, 33, and 35. After analyzing the relationship with prompt 29, it was shown that participants included both academic and non-academic factors in grades. Participants showed a strong agreement with the inclusion of the following non-academic factors when grading: extra credit, a teacher's personal judgment, assignments other than summative assessments, class participation, homework, turning in an assignment after its due date, and using grades to reward a student's high academic achievement. Participants showed a strong agreement with the inclusion of the following academic factors when grading: individual grades for group work, project-based assignments, the act of conferencing with a student about their grades throughout each grading period, and the belief that teachers should be accountable for following best-grading practices.

Synthesis of the strong positive relationship with prompt 30 included the following non-academic prompts: 9, 10, 11, 12, 18, 22, 25, and 27 and the following academic prompts: 29, 32, and 39. After analyzing the relationship with prompt 30, it was shown that participants included both academic and non-academic factors in grades. Participants showed a strong agreement with the inclusion of the following non-academic factors when grading: a student's ability level when grading, assignment completion, extra credit, a teacher's personal judgment, class participation, homework, turning in an assignment after its due date, and a student's effort. Participants showed a strong agreement with the inclusion of the following academic factors when grading: a student's progress towards achieving the course learning standards, project-based assignments, and district-provided professional development about best-grading practices.

Synthesis of the strong positive relationship with prompt 31 included the following non-academic prompts: 13, 16, and 28 and the following academic prompts: 33, 35, and 41. After analyzing the relationship with prompt 31, it was shown that participants both academic and non-academic factors in grades. Participants showed a strong agreement with the inclusion of the following non-academic factors when grading: assignments other than summative assessments, assignment completion, and using grades to reward a student's high academic achievement. Participants showed a strong agreement with the inclusion of the following academic factors when grading: project-based assignments, the act of conferencing with a student about their grades throughout each grading period, and the belief that teachers should be accountable for following best-grading practices.

Synthesis of the strong positive relationship with prompt 32 included the following non-academic prompts: 9, 11, 12, 18, 22, 25, and 28 and the following academic prompts: 29, 31, and 33. After analyzing the relationship with prompt 32, it was shown that participants included both academic and non-academic factors in grades. Participants showed a strong agreement with the inclusion of the following non-academic factors when grading: a student's ability level, extra credit, a teacher's personal judgment, class participation, homework, turning in an assignment after its due date, and using grades as a reward for a student's high academic achievement. Participants showed a strong agreement with the inclusion of the following academic factors when grading: a student's progress towards achieving the course learning standards, individual projects, and project-based assignments.

Synthesis of the strong positive relationship with prompt 33 included the following non-academic prompts: 13, 16, 18, 22, 27, and 28 and the following academic prompts: 29, 32, 35, and 41. After analyzing the relationship with prompt 33, it was shown that participants included both academic and non-academic factors in grades. Participants showed a strong agreement with the inclusion of the following non-academic factors when grading: assignments other than summative assessments, class participation, homework, a student's effort, and using grades to reward a student's high academic achievement. Participants showed a strong agreement with the inclusion of the following academic factors when grading: a student's progress towards achieving the course learning standards, individual projects, and project-based assignments, and the belief that teachers should be accountable for following best-grading practices.

Synthesis of the strong positive relationship with prompt 34 included academic prompt 31. After analyzing the relationship with prompt 34, it was shown that participants included an academic factor in grades. Participants showed a strong agreement with the inclusion of assigning grades for individual projects when grading.

Synthesis of the strong positive relationship with prompt 35 included the following non-academic prompts: 13, 15, 16, 27, and 28 and the following academic prompts: 29, 31, 33, and 41. After analyzing the relationship with prompt 35, it was shown that participants included both academic and non-academic factors in grades. Participants showed a strong agreement with the inclusion of the following non-academic factors when grading: assignments other than summative assessments, zeroes for missing assignments, assignment completion, a student's effort, and using grades as a reward for a student's high academic achievement. Participants showed a strong agreement with the

inclusion of the following academic factors when grading: a student's progress towards achieving the course learning standards, individual projects, project-based assignments, and the belief that teachers should be accountable for following best-grading practices.

Synthesis of the strong positive relationship with prompt 36 included non-academic prompts 10 and 23. After analyzing the relationship with prompt 36, it was shown that participants included non-academic factors in grades. Participants showed a strong agreement with the inclusion of the non-academic factors of assignment completion and note-taking when grading.

Synthesis of the strong positive relationship with prompt 37 included academic prompt 34. After analyzing the relationship with prompt 37, it was shown that participants included academic factors in grades. Participants showed a strong agreement with the inclusion of the academic factor of student portfolios when grading.

Synthesis of the strong positive relationship with prompt 39 included the following non-academic prompts: 11, 12, and 23 and academic prompt 30. After analyzing the relationship with prompt 39, it was shown that participants included both academic and non-academic factors in grades. Participants showed a strong agreement with the inclusion of the following non-academic factors when grading: extra credit, a teacher's personal judgment, and note-taking. Participants showed a strong agreement with the inclusion of the academic factor of individual grades for group work when grading.

Synthesis of the strong positive relationship with prompt 41 included the following non-academic prompts: 13, 16, 18, 27, and 28 and the following academic prompts: 29, 31, 33, and 35. After analyzing the relationship with prompt 41, it was

shown that participants included both academic and non-academic factors in grades.

Participants showed a strong agreement with the inclusion of the following non-academic

factors when grading: assignments other than summative assessments, assignment

completion, class participation, a student's effort, and using grades as a reward for a

student's high academic achievement. Participants showed a strong agreement with the

inclusion of the following academic factors when grading: a student's progress towards

achieving the course learning standards, individual projects, project-based assignments,

and conferencing with a student about their grades throughout each grading period.

     Synthesis of the strong positive relationship with prompt 43 included academic

prompts 49 and 55. After analyzing the relationship with prompt 43, it was shown that

participants included academic factors in grades. When grading, participants showed a

strong agreement with the inclusion of the academic factors of the belief that grades serve

as a reflection of student mastery of course learning standards and the belief that grades

are indicators of future performance on state exams.

     Synthesis of the strong positive relationship with prompt 44 included academic

prompts 45 and 48. After analyzing the relationship with prompt 44, it was shown that

participants included academic factors in grades. When grading, participants showed a

strong agreement with the inclusion of the academic factors of district-provided

collaboration time on early-release days and regular collaboration during common plan

time.

     Synthesis of the strong positive relationship with prompt 45 included the

following academic prompts: 44, 46, and 51. After analyzing the relationship with

prompt 45, it was shown that participants included academic factors in grades.

Participants showed a strong agreement with the inclusion of the following academic factors when grading: common plan time for teachers of the same content area to promote collaboration, district-provided collaboration time on late start days, and the availability of a field in their gradebook by which to indicate student mastery of course learning standards.

Synthesis of the strong positive relationship with prompt 46 included academic prompts 45 ad 51. After analyzing the relationship with prompt 46, it was shown that participants included academic factors in grades. When grading, participants showed a strong agreement with the inclusion of the academic factors of district-provided collaboration time on early release days and the availability of a field in their gradebook by which to indicate student mastery of course learning standards.

Synthesis of the strong positive relationship with prompt 47 included academic prompts 50 and 52. After analyzing the relationship with prompt 47, it was shown that participants included academic factors in grades. When grading, participants showed a strong agreement with the inclusion of the academic factors of allowing students to redo assignments and the belief that grades are intended to provide feedback to stakeholders about a student's mastery of the course learning standards.

Synthesis of the strong positive relationship with prompt 48 included academic prompts 44 and 51. After analyzing the relationship with prompt 48, it was shown that participants included academic factors in grades. When grading, participants showed a strong agreement with the inclusion of the academic factors of common plan time for teachers in the same content area to promote collaboration and the availability of a field in their gradebook by which to indicate student mastery of course learning standards.

Synthesis of the strong positive relationship with prompt 49 included academic prompts 43, 50, and 52. After analyzing the relationship with prompt 49, it was shown that participants included academic factors in grades. Participants showed a strong agreement with the inclusion of the following academic factors when grading: allowing students to redo assignments, the consideration of how their grading practices compare to other teachers' grading practices, and the belief that grades are intended to provide feedback to stakeholders about a student's mastery of the course learning standards.

Synthesis of the strong positive relationship with prompt 50 included academic prompts 47, 49, and 52. After analyzing the relationship with prompt 50, it was shown that participants included academic factors in grades. Participants showed a strong agreement with the inclusion of the following academic factors when grading: district-provided collaboration time on professional development days, allowing students to redo assignments, and the belief that grades are intended to serve as a reflection of a student's mastery of the course learning standards.

Synthesis of the strong positive relationship with prompt 51 included academic prompts 45, 46, and 48. After analyzing the relationship with prompt 51, it was shown that participants included academic factors in grades. Participants showed a strong agreement with the inclusion of the following academic factors when grading: district-provided collaboration time for teachers on early release days, district-provided collaboration time for teachers on late start days, and the practice of regularly collaborating with other grade level teachers during plan time.

Synthesis of the strong positive relationship with prompt 52 included academic prompts 47, 49, and 50. After analyzing the relationship with prompt 52, it was shown

that participants included academic factors in grades. Participants showed a strong agreement with the inclusion of the following academic factors when grading: district-provided collaboration time for teachers on professional development days, the belief that grades are intended to serve as a reflection of a student's mastery of the course learning standards, and the belief that grades are intended to provide feedback to stakeholders about a student's mastery of the course learning standards.

Synthesis of the strong positive relationship with prompt 53 included academic prompt 56. After analyzing the relationship with prompt 53, it was shown that participants included an academic factor in grades. Participants showed a strong agreement with the inclusion of the academic factor of weighing summative assessments heavier than formative assessments when grading.

Synthesis of the strong positive relationship with prompt 55 included academic prompt 43. After analyzing the relationship with prompt 55, it was shown that participants included an academic factor in grades. Participants showed a strong agreement with the inclusion of the academic factor of comparing grading practices with other teachers' grading practices in the same content area when grading.

Synthesis of the strong positive relationship with prompt 56 included academic prompt 53. After analyzing the relationship with prompt 56, it was shown that participants included an academic factor in grades. Participants showed a strong agreement with the inclusion of the academic factor of allowing students to retake a test they failed when grading.

Strong Inverse Relationships

As previously stated, for prompts 6 through 41 and 43 through 56, potential relationships between participant responses were analyzed with the application of the PPMCC. Analysis on responses to the prompts represented by strongly disagree, disagree, neither agree nor disagree, somewhat agree, and strongly agree yielded significant relationships (see Appendix N). This report discusses the relationships found to be strong and inversely related, with an *r* value less than -0.750.

Strong inverse relationships were found between a number of prompts in the study survey, as well. Prompts are grouped together according to strong agreement in response rates, among the topics presented in the survey (see Table 13). For example, participants responded with significant disagreement to prompt 6. Responses to prompts 15, and 41 indicated the opposite trends of agreement as prompt 6; when strongly disagree was the strongest percentage in prompt 6, strongly agree was the strongest percentage in prompts 15 and 41.

Table 13

Strong Inverse Relationships in Survey Response Rates

Q	Strong Inverse Relationship (< -0.750) with
# 6	15, 41
# 7	15
# 8	15
# 9	
# 10	
# 11	
# 12	
# 13	37
# 14	15, 41
# 15	6, 7, 8, 14, 17, 19, 20, 21, 24, 26,34, 37
# 16	
# 17	15, 41

Table 13. Continued.

# 18	
# 19	15
# 20	15, 41
# 21	15
# 22	37
# 23	
# 24	15
# 25	
# 26	15
# 27	
# 28	37
# 29	
# 30	
# 31	
# 32	
# 33	37
# 34	15, 37
# 35	34, 37
# 36	
# 37	13, 15, 22, 27, 33, 35
# 38	
# 39	
# 40	
# 41	6, 14, 17, 20
# 42	
# 43	
# 44	53
# 45	
# 46	
# 47	54
# 48	
# 49	
# 50	
# 51	
# 52	53
# 53	44
# 54	47, 52
# 55	
# 56	

Synthesis of the strong inverse relationship with prompt 6 included non-academic prompt 15 and academic prompt 41. After analyzing the relationship with prompt 6, it was shown that when participants strongly disagreed that they considered a student's behavior in grades, they strongly agreed with the inclusion of a non-academic factor and an academic factor in grades. Participants showed strong agreement with the inclusion of the non-academic factor of zeroes for missing assignments in grades. Participants also showed strong agreement with the inclusion of the academic factor of the belief that teachers should be accountable for following best-grading practices in grades.

Synthesis of the strong inverse relationship with prompt 7 included non-academic prompt 15. After analyzing the relationship with prompt 7, it was shown that when participants strongly disagreed that grades are a means to punish students for low academic achievement, they strongly agreed with the inclusion of the non-academic factor of zeroes for missing assignments in grades.

Synthesis of the strong inverse relationship with prompt 8 included non-academic prompt 15. After analyzing the relationship with prompt 8, it was shown that when participants showed strong disagreement that they assign points for bringing daily supplies to class, they showed strong agreement with the inclusion of the non-academic factor of zeroes for missing assignments in grades.

Synthesis of the strong inverse relationship with prompt 13 included academic prompt 37. After analyzing the relationship with prompt 13, it was shown that when participants strongly agreed that they assign grades for assignments other than assessments, they strongly disagreed about the inclusion of the academic factor of

requiring students to reflect on their achievement of course learning standards in a journal

or notebook in grades.

Synthesis of the strong inverse relationship with prompt 14 included non-

academic prompt 15 and academic prompt 41. After analyzing the relationship with

prompt 14, it was shown that when participants strongly agreed with considering a

student's involvement in extracurricular activities, they strongly agreed about the

inclusion of a non-academic factor and an academic factor in grades. Participants showed

strong agreement of the inclusion of the non-academic factor of assigning students zeroes

for missing assignments in grades. Participants also showed strong agreement with the

academic factor of the inclusion of the belief that teachers should be accountable for

following best-grading practices in grades.

Synthesis of the strong inverse relationship with prompt 15 included non-

academic prompts 6, 7, 8, 14, 17, 19, 20, 21, 24, 26 and academic prompts 34 and 37.

After analyzing the relationship with prompt 15, it was shown that when participants

strongly agreed that they assign zeroes for missing assignments, they showed strong

disagreement about the inclusion of non-academic and academic factors in grades.

Participants showed strong disagreement with the inclusion of the following non-

academic factors in grades: a student's behavior, using grades as a means to punish

students for low academic achievement, assigning points for bringing daily supplies to

class, a student's involvement in extracurricular activities, a student's GPA, a student's

gender, a student's physical appearance, the aesthetic appearance of a student's work, and

a student's personality. Participants also showed strong disagreement with the inclusion

of the academic factors of student portfolios and the act of requiring students to reflect on their achievement of course learning standards in a journal or notebook.

Synthesis of the strong inverse relationship with prompt 17 included non-academic prompt 15 and academic prompt 41. After analyzing the relationship with prompt 17, it was shown that when participants disagreed that they considered a student's GPA in grades, they agreed with the inclusion of a non-academic factor and an academic factor in grades. Participants showed strong agreement with the inclusion of the non-academic factor of zeroes for missing assignments in grades. Participants also showed strong agreement with the inclusion of the academic factor of the belief that teachers should be accountable for following best-grading practices in grades.

Synthesis of the strong inverse relationship with prompt 19 included non-academic prompt 15. After analyzing the relationship with prompt 19, it was shown that when participants disagreed that they consider a student's gender when grading, they agreed with the inclusion of the non-academic factor of zeroes for missing assignments in grades.

Synthesis of the strong inverse relationship with prompt 20 included non-academic prompt 15 and academic prompt 41. After analyzing the relationship with prompt 20, it was shown that when participants disagreed that they have lowered a student's grade for demonstrating poor behavior, they agreed with the inclusion of a non-academic factor and an academic factor in grades. Participants showed strong agreement with the inclusion of the non-academic factor of zeroes for missing assignments in grades. Participants also showed strong agreement with the inclusion of the academic

factor of the belief that teachers should be accountable for following best-grading practices in grades.

Synthesis of the strong inverse relationship with prompt 21 included non-academic prompt 15. After analyzing the relationship with prompt 21, it was shown that when participants disagreed that they consider the physical appearance of a student when grading, they agreed with the inclusion of the non-academic factor of zeroes for missing assignments in grades.

Synthesis of the strong inverse relationship with prompt 22 included academic prompt 37. After analyzing the relationship with prompt 22, it was shown that when participants agreed that they assign grades for homework, they disagreed with the inclusion of the academic factor of requiring students to reflect on their achievement of the course learning standards in a journal or notebook.

Synthesis of the strong inverse relationship with prompt 24 included non-academic prompt 15. After analyzing the relationship with prompt 24, it was shown that when participants disagreed that they consider the aesthetic appearance of the student's work when grading, they agreed with the inclusion of the non-academic factor of zeroes for missing assignments in grades.

Synthesis of the strong inverse relationship with prompt 26 included non-academic prompt 15. After analyzing the relationship with prompt 26, it was shown that when participants disagreed that they consider a student's personality when grading, they agreed with the inclusion of the non-academic factor of zeroes for missing assignments in grades.

Synthesis of the strong inverse relationship with prompt 28 included academic prompt 37. After analyzing the relationship with prompt 28, it was shown that when participants agreed that grades should reward students for high academic achievement, they disagreed with the inclusion of the academic factor of requiring students to reflect on their achievement of course learning standards in a journal or notebook in grades.

Synthesis of the strong inverse relationship with prompt 33 included academic prompt 37. After analyzing the relationship with prompt 33, it was shown that when participants agreed that they assign grades for project-based assignments, they disagreed with the inclusion of the academic factor of requiring students to reflect on their achievement of course learning standards in a journal or notebook in grades.

Synthesis of the strong inverse relationship with prompt 34 included non-academic prompt 15 and academic prompt 37. After analyzing the relationship with prompt 34, it was shown that when participants disagreed that they include student portfolios in grades, they agreed with the inclusion a non-academic factor and an academic factor in grades. Participants showed strong agreement with the inclusion of the non-academic factor of zeroes for missing assignments. Participants also showed strong agreement with the inclusion of the academic factor of requiring students to reflect on their achievement of course learning standards in a journal or notebook in grades.

Synthesis of the strong inverse relationship with prompt 35 included academic prompts 34 and 37. After analyzing the relationship with prompt 35, it was shown that when participants strongly agreed that they conference with students about their grades throughout each grading period, they strongly disagreed with the inclusion of academic factors in grades. Participants showed strong disagreement with the inclusion of the

academic factors of student portfolios and requiring students to reflect on their

achievement of course learning standards in a journal or notebook in grades.

Synthesis of the strong inverse relationship with prompt 37 included non-

academic prompts 13, 15, 22, and 27 and academic prompts 33 and 35. After analyzing

the relationship with prompt 37, it was shown that when participants disagreed that they

include a requirement that students reflect on their achievement of course learning

standards in a journal or notebook in grades, they agreed with the inclusion of non-

academic and academic factors in grades. Participants showed strong agreement with the

inclusion of the following non-academic factors: grades for assignments other than

assessments, zeroes for missing assignments, grades for homework, and a student's

effort. Participants also showed strong agreement with the inclusion of the academic

factors of project-based grading and requiring students to reflect on their achievement of

course learning standards in a journal or notebook in grades.

Synthesis of the strong inverse relationship with prompt 41 included non-

academic prompts 6, 14, 17, and 20. After analyzing the relationship with prompt 41, it

was shown that when participants agreed with the belief that teachers should be

accountable for following best-grading practices, they agreed with the inclusion of non-

academic factors in grades. Participants showed strong agreement with the inclusion of

the following non-academic factors: a student's behavior, a student's involvement in

extracurricular activities, and a student's GPA in grades.

Synthesis of the strong inverse relationship with prompt 44 included academic

prompt 53. After analyzing the relationship with prompt 44, it was shown that when

participants disagreed that their district provides common planning time for teachers in

the same content area to promote collaboration, they agreed with the inclusion of the academic factor of allowing a student to re-take a test if they fail it in grades.

Synthesis of the strong inverse relationship with prompt 47 included academic prompt 54. After analyzing the relationship with prompt 47, it was shown that when participants agreed that their district provides collaboration time for teachers on professional development days, they disagreed with the inclusion of the academic factor of summative, and not formative, assessments it in grades.

Synthesis of the strong inverse relationship with prompt 52 included academic prompt 53. After analyzing the relationship with prompt 52, it was shown that when participants agreed that they allow students to re-do assignments, they agreed with the inclusion of the academic factor of allowing a student to re-take a test if they fail it in grades.

Synthesis of the strong inverse relationship with prompt 53 included academic prompt 44. After analyzing the relationship with prompt 53, it was shown that when participants agreed that they allow students to re-take a test if they fail it, they disagreed that their district provides common planning time for teachers in the same content area to promote collaboration.

Synthesis of the strong inverse relationship with prompt 54 included academic prompts 47 and 52. After analyzing the relationship with prompt 54, it was shown that when participants disagreed that they grade summative assessments and not formative assessments, they agreed with the inclusion of the academic factors of allowing students to red-do assignments in grades and that their district provides collaboration time for teachers on professional development days.

Research Question One: What types of support work best for teachers trying to implement best-grading practices?

Participants identified a variety of supports that work best including collaboration with peers and administrators, pre-service training, and professional development. Districts were able to reinforce and expand upon these supports in attempt to promote implementation of best-grading practices by their teachers. Participants indicated that the availability of these district-provided supports, in particular, increased their likelihood of implementing best-grading practices.

Approximately 71% of participants indicated that they collaborated with their peers within the range of once per day to once per month. Participants stated that they discussed academic topics including pacing, instructional strategies, vertical alignment, course learning standards, curriculum, assessments, grading practices, student achievement, and modifications and accommodations for special education students. They also indicated that they discussed non-academic topics consisting of classroom management, student discipline, and facilities management. In response to the survey, one participant stated, "I wish we were able to collaborate with peers more. There are so many things occurring in our classrooms that we should be discussing on a regular basis."

Only about 30% of participants indicated that they received training on best-grading practices in their undergraduate, pre-service coursework. Teacher #1 stated,

> There was very little grading practice training in my college. What training I did receive was theoretical. We didn't ever actually practice grading or anything like that. I think most teachers' grading practices have developed as a result of their student teaching, classroom experience, and mentors and peers. I know mine did.

Similarly, Teacher #3 stated,

> I sometimes use expert recommended grading practices to inform me, but I am the sole decider about what should be implemented in my classroom. This is a teacher's right because I know best about what works for my class. Only principals should be able to override that.

Despite an overall lack of pre-service training in best-grading practices, approximately 68% of participants stated that they believed teachers should be accountable for following best-grading practices. This is further complicated by the fact that only about 44% of participants indicated that their districts provided professional development about best-grading practices. However, several participants stated that they felt they could benefit from their districts offering more professional development about best-grading practices.

Table 14 indicated that 48% of participants responded that their districts provided time for teachers in the same content area to meet to align grading practices whereas 33% stated that their districts had not provided time for such.

Table 14

District-Provided Content Area Teacher Collaboration Time

#	Answer	%	Count
1	Strongly disagree	23.53%	8
2	Somewhat disagree	8.82%	3
3	Neither agree nor disagree	20.59%	7
4	Somewhat agree	23.53%	8
5	Strongly agree	23.53%	8
	Total	100%	34

Table 15 indicated that 44% of participants stated that their districts provided

professional development about best-grading practices and 39% stated that their districts

had not.

Table 15

District-Provided Professional Development About Best-Grading Practices

#	Answer	%	Count
1	Strongly disagree	23.53%	8
2	Somewhat disagree	14.71%	5
3	Neither agree nor disagree	17.65%	6
4	Somewhat agree	26.47%	9
5	Strongly agree	17.65%	6
	Total	100%	34

Table 16 showed that 52% of teachers stated their administrators spoke with them

within the range of once per quarter to once per semester.

Table 16

Administrator-Teacher Grading Practice Dialogue Frequency

#	Answer	%	Count
1	Never	11.76%	4
2	1 x daily	0.00%	0
3	1 x weekly	5.88%	2
4	1 x monthly	14.71%	5
5	1 x quarterly	26.47%	9
6	1 x per semester	26.47%	9
7	1 x yearly	14.71%	5
	Total	100%	34

Research Question Two. Which academic factors do teachers use when grading?

Participants indicated that they included the following academic factors when grading: Student progress towards achieving the course learning goals (76%), project-based assignments (74%), with individual grading for such (65%), individual projects (85%), individual grades for group work (76%), summative assessments (50%), and assignments that have been re-done (71%). Participants further indicated that they considered the following beliefs and practices pertaining to academic factors when grading: self-reflection via comparison with colleagues in the same content area about grading practices (56%), conferences with students about their grades throughout each grading period (79%), teachers should be accountable for following best-grading practices (68%), grades should be a direct reflection of student master of the course learning standards (88%), grades are a means by which to provide feedback to stakeholders about student progress of the course learning standards (74%), and grades are an indication of how students should perform on state exams (56%).

Both survey and interview participants noted that they also considered a student's IEP/modifications, as well, when grading. For example, Teacher #3 stated, "I typically only allow general education students to re-take tests at the end of each quarter. However, if a student had an IEP I would allow them re-takes at any point in time."

While 56% of survey participants indicated that they self-reflected by comparing their grading practices with peers in the same content area, all interview participants communicated that they felt teachers, in general, could benefit from more district-provided time in which they could collaborate with peers about grading practices. Teacher #3 stated,

It would be nice to have more district-provided workshops or professional

development days about grading practices and then have more district-provided

time to actually strategize how what we learned could be best implemented. I

don't have a very high opinion of a lot of other teacher's grading practices. I think

that if more teachers were better-educated by their districts about grading

practices, there might be more teachers using grading practices I would want to

potentially emulate.

Research Question Three. Why do teachers choose to use academic factors

when grading?

Participants stated that they used academic factors when grading because they

believed grades included several measurable academic factors that served to indicate a

student's academic aptitude. About 88% of survey participants indicated that they believe

that grades should be a direct reflection of student mastery of the course learning

standards. Seventy-four percent stated that they believed that grades should provide

feedback about student progress of the course learning standards to stakeholders. Fifty-

six percent of participants indicated that they believed that grades served as an indication

of how students should be expected to perform on state exams. As stated by Teacher #1,

"Grades should demonstrate a comprehensive picture of a student's abilities to show their

strengths and weaknesses so that teachers know students are on track to achieve the

course learning objectives."

Teacher #3 further stated that she felt most teachers liked to include academic

factors when grading because they are less disputable than non-academic factors. She

stated that when grades were not primarily based on academic factors, "you end up

giving A's to the students who are the best liars or, rather, the students who are able to come up with the best excuses."

Research Question Four. Which non-academic factors do teachers use when grading?

Participants indicated that they included the following non-academic factors when grading: Student effort (71%), student ability level (59%), extra credit (68%), zeroes for missing assignments (50%), formative assessments (71%), assignment completion (94%), class participation (71%), homework (82%), the act of using grades as a reward for high academic achievement (74%), and student responsibility (62%). Additionally, some survey participants further noted that they included attendance, a student's personal level of interest in the subject, behavior, and a student's home environment when grading.

All six teachers interviewed emphasized consideration of factors associated with the concepts of student responsibility and effort, when grading. Teacher #3 stated,

> Outside of IEP's, general education students who take the time to put forth more effort than their peers should be rewarded academically. When I'm grading, I try not to see whose work it is. But, if while I'm grading I see a small error the student has made, I will then look at the name to see if I know of things that particular student struggles with and take that into consideration.

Teacher #1 also expressed a similar sentiment in regard to why she believes that teachers should be able to give students zeroes that are included in their grades, "No work. No points. No excuses. That's life."

Research Question Five. Why do teachers choose to use non-academic factors when grading?

Behavioral related non-academic factors used to aid in classroom management were emphasized by participants as integral to student learning. For example, teachers indicated that student behavior, responsibility, and zeroes were essential to maintaining motivated students in an optimal learning environment. Several participants indicated that excluding these factors from their grading practices would result in decreased student achievement, because teachers would spend more time on discipline and less time teaching.

Teacher #1 stated,

Behavior should be included in grades. If it isn't, it punishes the teacher. Otherwise there's no big consequence for a student doing something like sleeping in class. If they sleep in class, why should I have to take more class time to re-teach them or when they chose not to engage in learning?

She continued,

And meeting deadlines is important. It's an important part of growing up. In the workforce, you can't turn in stuff late. You get fired for doing things like that. Students either learn to turn stuff in on time or they incur the penalty - a 20% deduction in their score per day it's late. As teachers, it's our job to teach students more than just academics. We spend an awful lot of time with these kids to teach them nothing else, and to not reflect that in their grade, doesn't make sense.

Teacher #2 agreed,

I don't see separating academic and non-academic factors in grades as effective because then teachers would have to suffer students being able to turn in

everything all at once at the very end of the quarter or semester. Grades should be used to motivate students.

H2. There are no differences in perception of best-grading practices, measured by a Likert-scale survey, among the following support structures: Frequency of teacher's discussions with peers regarding best-grading practices, frequency of teacher's discussions with building administrators regarding best practices, and frequency of the teacher's engagement in self-reflection in which the teacher is engaged.

Teacher #1:

Different administrators have different opinions on grading practices and it makes it hard for teachers and students to navigate that. I teach both middle and high school students. But the middle school and high school principals have very different opinions about fairness in grades. This results in some students' receiving an unfair academic advantage compared to other students where the principal has different expectations for their teachers about grading practices. And there's a burden placed on the teacher to have to be shifting from one expectation to another depending on which administrator they're dealing with at that particular moment.

Teacher #2 agreed,

Teachers always have to consider the grading practices that their principal believes in. And a lot of the time, those preferences are based on the principal's experiences when they were a teacher that they have carried over to their beliefs as an administrator. Principals set the standards for what they want to see from their teachers in terms of grading practices because the principals are the ones

evaluating the teachers.

Conclusion

The results of this study added to the body of research about teachers' grading practices and the implementation of expert recommended grading practices. There were many factors that contributed to teachers' grading practices. These were not limited to only academic and non-academic factors, but, rather they also included factors, such as teachers' overall beliefs and values about grading and their personal experiences and training, as well as that of their administrators.

As stated by Santoro (2011), "Teaching is an intellectual and moral practice fraught with contradictions, impediments, and challenges both quotidian and extraordinary" (p. 1). Those called to the teaching profession are people who were typically motivated by the moral rewards it provided- being a public servant, bettering the future, and crafting knowledge (Santoro, 2011). However, then-current school reform threatened the existence of the moral rewards of teaching. Therefore, researchers found that school reform resulted in an international demoralization of the teaching profession (Nodding, 2008; Santoro, 2011; Tsang & Liu, 2016; Wang, 2013). Santoro (2011) defined demoralization as "consistent and persistent frustrations in accessing the moral rewards of teaching" wherein teachers feel they "can no longer do good work or teach 'right'" (p. 2). Distinct from the concept of teacher burnout, teacher demoralization was believed to have occurred as the result of societal pressures often associated with the organizational hierarchy that existed within schools (Santoro, 2012). Therefore, this demoralization would also include teachers' grading practices.

Data obtained from both the survey and interviews revealed that the majority of participants believed teachers should be able to exercise a degree of personal judgment in

assigning grades. Participants stated this privilege was important to them because it validated their years of experience and expertise in teaching. Data indicated that 71% of survey participants and 100% of interview participants stated that they believed teachers should be able to exercise their judgment when grading; thereby, exercising a moral reward of the profession. This sentiment was further supported by the fact that 61% of survey participants indicated that they self-reflected about their grading practices within the range of once daily to once weekly. Therefore, data showed the practice of grading and teachers' ability to exercise personal judgment when grading were both important to teachers and were frequently on their minds.

All interviewed participants felt it would negatively impact their grading practices, as well as their teaching, if they did not feel their administrators had confidence in their ability to assign grades as they deemed just. These findings were consistent with research that showed that teachers did not perform as well when they were subjected to strict supervisory support services. Researchers found that teachers at schools with high morale and those at schools with low morale had differing support systems by which they and their work were supervised. When teachers felt that their supervision was strict, they tended to develop feelings of dissatisfaction (and this negatively impacted their grading practices) (Ingersoll, 2003).

The survey showed that approximately 62% of participants engaged in dialogue with their administrators about grading practices within the range of once per day to once per week. However, while participants reported routine dialoging about grading practices, they further stated that administrator opinions about grading practices and the enforcement of grading policies considerably differed. According to several participants,

this sometimes made teachers feel as though administrators' expectations about teachers'

grading practices were inconsistent and could therefore be difficult to comply with.

      The findings of this study provide perspective about teachers' demographics, the

academic and non-academic factors teachers included in their grading practices, the

reasons why teachers included academic and non-academic factors in their grading

practices, and the types of support that worked best for districts when attempting to

implement best-grading practices. The results of this study suggested that educational

leaders should focus resources on improving the quality of teacher preparation programs'

grading practice instruction, and professional development about best-grading practices.

Investments in supports, such as collaborative time for teachers, teacher and

administrative grading practice training, and written district grading policies may be of

particular benefit to schools hoping to implement best-grading practices with fidelity.

Chapter Five provides a summary of the study, findings, reflections, and

recommendations.

Chapter Five: Discussion and Reflection

Introduction

Chapter Five includes a summary of the study and a discussion of the findings. A brief review of the methodology is also provided. Reflections and recommendations related to the findings are presented, as well.

This mixed-methods phenomenological study examined a sample of grade 7 through 12 core content area teachers at four Missouri school districts. The main purpose of this study was to investigate possible patterns linking demographic characteristics of teachers and grading practices. These practices were cross-referenced to the level of implementation of best-grading practices, as it pertained to the amount of time and support systems a district provided for its teachers. The desire to improve teachers' grading practices in an attempt to ensure fair and accurate grading for students was the motivation behind this study.

The following five research questions guided the study.

RQ1: What types of support work best for teachers trying to implement best-grading practices?

RQ2: Which academic factors do teachers use when grading?

RQ3: Why do teachers choose to use academic factors when grading?

RQ4: Which non-academic factors do teachers use when grading?

RQ5: Why do teachers choose to use non-academic factors when grading?

In addition, the following hypotheses were tested:

H1: There will be differences in demographic characteristics within respondents to a Likert-scale survey, with regard to best-grading practices, among the following

teacher-related variables: Teachers' age, teachers' gender, teachers' levels of education, teacher's years of service, and teachers' content areas.

H2: There will be differences in perception of best-grading practices, measured by a Likert-scale survey, among the following support structures: Frequency of teachers' discussions with peers regarding best-grading practices, frequency of teachers' discussions with building administrators regarding best practices, and frequency of the teachers' engagement in self-reflections in which the teacher is engaged.

H3: There will be differences in perception of agreement topic-to-topic, and agreement-to-disagreement with survey statements, with regard to best-grading practices, measured by a Likert-scale, with regard to the following support characteristics: A student's behavior, grades as a means of punishment, student responsibility, a student's ability level, assignment completion points, extra credit, a teacher's personal judgment, assignments other than assessments, a student's involvement in extracurricular activities, zeroes for missing assignments, a student's GPA, class participation, a student's gender, a student's physical appearance, homework, note-taking, aesthetic appearance of a student's work, a student's personality, student effort, using grades as a reward, student progress towards achieving the course learning standards, individual grades for group work, individual projects, project-based grading, project-based assignments, student portfolios, conferences with students about individual grades throughout each grading period, student-set personal learning goals for the course, student personal reflection on achievement of the course learning standards, district-provided time for teachers in the same content area to meet to align grading practices, district-provided professional development about best-grading practices, pre-service training for teachers on best-

grading practices, the belief that teachers should be accountable for following best-grading practices, teacher-to-teacher grading practice comparisons, district-provided common planning time for teachers in the same content area, district-provided collaboration time for teachers on early release days, district-provided collaboration time for teachers on late start days, district-provided collaboration time for teachers on professional development days, teacher collaboration, the belief that grades should be a direct reflection of student mastery of course learning standards, grades as a means to provide feedback to stakeholders, gradebook field that indicates student mastery of learning competencies, re-do assignments, test re-takes, and summative and formative assessments.

H4: There will be relationships among levels of agreement topic-to-topic, with regard to best-grading practices, measured by a Likert-scale survey, with regard to the following support characteristics: A student's behavior, grades as a means of punishment, student responsibility, a student's ability level, assignment completion points, extra credit, a teacher's personal judgment, assignments other than assessments, a student's involvement in extracurricular activities, zeroes for missing assignments, a student's GPA, class participation, a student's gender, a student's physical appearance, homework, note-taking, the aesthetic appearance of a student's work, a student's personality, student effort, using grades as a reward, student progress towards achieving the course learning standards, individual grades for group work, individual projects, project-based grading, project-based assignments, student portfolios, conferences with students about individual grades throughout each grading period, student-set personal learning goals for the course, student personal reflection on achievement of the course learning standards, district-

provided time for teachers in the same content area to meet to align grading practices, district-provided professional development about best-grading practices, pre-service training for teachers on best-grading practices, the belief that teachers should be accountable for following best-grading practices, teacher-to-teacher grading practice comparisons, district-provided common planning time for teachers in the same content area, district-provided collaboration time for teachers on early release days, district-provided collaboration time for teachers on late start days, district-provided collaboration time for teachers on professional development days, teacher collaboration, the belief that grades should be a direct reflection of student mastery of course learning standards, grades as a means to provide feedback to stakeholders, gradebook field that indicates student mastery of learning competencies, re-do assignments, test re-takes, and summative and formative assessments. These questions sought to provide perspective about teachers' demographics, the academic and non-academic factors teachers included in their grading practices, the reasons why teachers included academic and non-academic factors in their grading practices, and the types of support that worked best for districts when attempting to implement best-grading practices. For these five questions, both qualitative and quantitative methods generated data used to identify patterns relating to teachers' demographics and their grading practices, as well as the types of support districts could use to support the implementation of best-grading practices.

The findings of this study supported the two sub-hypotheses for H1. First, the sub-hypothesis that there is a difference in age representation among participants for two comparison age-group pairings was supported; but, a third age-group pairing was not

supported. Second, the sub-hypothesis that there is a difference in the percentage of teachers earning degrees higher than a bachelor's was also not supported.

However, the findings for H1 resulted in three sub-hypotheses that were not supported. First, the sub-hypothesis that there is a difference in gender representation was not supported. Second, the sub-hypothesis that there is a difference in representation among content areas was also not supported.

To check for significant differences in the frequency of three different types of collaboration for H2, an ANOVA was applied to data for prompts 43, 58, and 60. Considering administrative dialogue, peer collaboration, and self-reflection as types of collaboration, with regard to the sub-hypothesis and that there were differences in the types of collaboration utilized; the researcher did not support the hypothesis. Additionally, in considering the frequency of self-reflection in use of grading practices, with regard to the sub-hypothesis that there were differences in the frequency of self-reflection, the researcher did not support the hypothesis.

To check for differences in agreement on the topics in the survey within the participant group, the researcher considered Hypothesis 3 - There were differences in perception of agreement topic-to-topic, and agreement-to-disagreement with survey statements, with regard to best-grading practices, measured by a Likert-scale, with regard to the following support characteristics: A student's behavior, grades as a means of punishment, student responsibility, a student's ability level, assignment completion points, extra credit, a teacher's personal judgment, assignments other than assessments, a student's involvement in extracurricular activities, zeroes for missing assignments, a student's GPA class participation, a student's gender, a student's physical appearance,

homework, note-taking, aesthetic appearance of a student's work, a student's personality,

student effort, using grades as a reward, student progress towards achieving the course

learning standards, individual grades for group work, individual projects, project-based

grading, project-based assignments, student portfolios, conferences with students about

individual grades throughout each grading period, student-set personal learning goals for

the course, student personal reflection on achievement of the course learning standards,

district-provided time for teachers in the same content area to meet to align grading

practices, district-provided professional development about best-grading practices, pre-

service training for teachers on best-grading practices, the belief that teachers should be

accountable for following best-grading practices, teacher-to-teacher grading practice

comparisons, district-provided common planning time for teachers in the same content

area, district-provided collaboration time for teachers on early release days, district-

provided collaboration time for teachers on late start days, district-provided collaboration

time for teachers on professional development days, teacher collaboration, the belief that

grades should be a direct reflection of student mastery of course learning standards,

grades as a means to provide feedback to stakeholders, gradebook field that indicates

student mastery of learning competencies, re-do assignments, test re-takes, and

summative and formative assessments. An ANOVA was applied to check for differences

in agreement on the topics listed in Hypothesis 3. No differences were established;

therefore, no individual t-tests or z-tests were applied.

Concerning the analysis for agreement versus disagreement topic-to-topic on

survey prompts 6 through 56, no significant differences were found when ANOVA was

applied. The researcher considered the sub-hypothesis, there were differences in

perceptions, with regard to survey prompts, when considering topic-to-topic discussion. The researcher did not support the hypothesis.

Also for H3, analysis of responses to survey prompts 6 through 41 and 43 through 56 was completed by applying a *z*-test for difference in proportions to the percentage of agreement and percentage of disagreement to each prompt. ANOVA was applied to the percentage of agreement versus the percentage of disagreement for survey prompts 6 through 41, question-to-question to identify whether strong differences in perception existed. Significant differences were not found. With regard to the sub-hypothesis, there were differences in perceptions, with regard to survey prompts, when considering topic-to-topic discussion, the researcher did not support the hypothesis.

Concerning the analysis for agreement versus disagreement difference topic-to-topic on survey prompts 43 through 56, no significant differences were found when ANOVA was applied. The researcher considered the sub-hypothesis, there were differences in perceptions, with regard to survey prompts, when considering topic-to-topic discussion. The researcher did not support the hypothesis. No significant differences were established, question-to-question. Therefore, no individual *z*-tests for difference in proportion were applied, with regard to topic-to-topic, or question-to-question discussion.

To apply Hypothesis 3 to perception topics represented by individual questions, a *z*-test for difference in proportions was applied to check for differences in agreement versus disagreement on the topics covered by individual questions and listed in Hypothesis 3. The following question prompts indicated significant disagreement in participant response rates: 6, 7, 8, 14, 17, 20, 24, 37, 44, 45, 46, and 51. Complete

disagreement was indicated by participants for prompts 19, 21, and 26. Significant

agreement was indicated by participants for the following prompts: 9, 11, 12, 13, 16, 18,

22, 25, 27, 28, 29, 30, 31, 32, 33, 35, 36, and 41, 43, 47, 49, 50, 52, 53, 55, and 56. There

were no significant differences found for prompts 10, 15, 23, 34, 38, 39, and 40, 48, and

54. Though trends of strong agreement were found in responses that allowed a grouping

of prompts that received similar response patterns from participants, there were no

significant differences found in a comparison of the level of response rates when

considered topic-to-topic.

For H4, prompts 6 through 41 and 43 through 56 were analyzed with the

application of the PPMCC to determine potential relationships between participant

responses. Strong positive relationships were found between a number of prompts in the

study survey. Strong inverse relationships were found between a number of prompts in

the study survey, as well. Prompts with strong positive relationships were grouped

together according to strong agreement in response rates among the topics presented in

the survey (see Table 11) and for strong inverse relationships (see Table 13).

Triangulation of Results

Utilizing a mixed methods approach, two data sets were collected in this study to

address the five research questions. Qualitative and quantitative data were obtained

through an electronic survey of teachers that consisted of 60 questions. Qualitative data

were also obtained from semi-structured interviews with teachers.

There were two phases of data collection. In the first phase, data were collected

through an electronic survey to gain a better understanding of teacher demographics, the

factors included in their grading practices, and the types of support their districts

provided. In the second phase, data were collected through interviews with teachers to gain a better understanding of the reasons why they included specific factors in their grading practices, their beliefs about grading practices, and the obstacles and recommendations for districts in regard to the types of support that best facilitated the implementation of best-grading practices.

The data sets were designed to aide the researcher in her investigation of the relationships between teacher demographic characteristics, grading, and the types of support provided by districts. The researcher cross-referenced these practices to the level of implementation of best-grading practices as it pertained to the amount of time and types of support a district provided. Specific demographics explored included: gender, age, content area, years of service and educational level. Data collected from the survey and interview were presented in Chapters Three and Four.

This study included two levels of participation. Superintendents at the four participating school districts were asked to forward the electronic survey to all grade 7 through 12 core content area teachers in their respective districts, inviting them to participate in the study. Thirty-four teachers completed the survey. Additionally, each superintendent was asked to forward the names of teachers they felt would be ideal candidates to interview. Four teachers participated in an interview.

Hypothesis and Research Questions

After analyzing the data, H1 was supported for two comparison pairings of age groups, but was not supported for a third pairing. There were significant differences between the number of 30-somethings compared to the number of 20-somethings and between the number of 40-somethings compared to the number of 30-somethings. The

sub-hypothesis that there is a difference in the percentage of teachers earning degrees higher than a bachelor's was also supported. The number of teachers with a master's degree was observably higher, and also significantly higher than, the number of teachers with a Bachelor's degree.

Therefore, approximately 79% of teachers were within the range of 30 to 49 years old. Whereas, about 6% of teachers were 20 to 29 years old, and about 15% were 50 to 99 years old. Given that more participants possessed a master's degree than simply a bachelor's degree alone, the findings reflected that there was a consistent gap in collegiate grading practice training- even at the graduate school level. This implied that the majority of participants were veteran teachers who, despite having completed graduate level coursework, had to rely on their personal experiences and that of their peers' to develop their grading practices as opposed to receiving formal training on such by a collegiate institution. This deficiency in training was apparent as a finding in this study, as well as other studies reported in expert literature.

Also in regards to H1, the sub-hypothesis that there is a difference in gender representation was not supported. Though observably more female teachers were employed, the findings were not significantly different and were not significantly more than the other gender. The sub-hypothesis, that there is a difference in the number of years of teaching experience, was also not supported. Though the number of years of experience was observably more frequent on the higher end of the range, the differences in years of experience throughout the data were not significant. Finally, the sub hypothesis that there is a difference in representation among content areas, was also not supported. There were no significant differences in the categories of content area (ELA,

S.S., Science, and Math) represented among the participants; therefore, no one content area was influencing the outcomes of the perceptions measured for this study.

After analyzing the data, H2 was not supported because there were no significant differences found in the types of collaboration (administrative dialogue, peer collaboration, and self-reflection) in which teachers participated. In addition, there were no significant differences in the types of collaboration in which teachers participated. Finally, there were no significant differences in the frequency with which teachers applied self-reflection to their own grading practices. Approximately 73% of survey participants reported that they engaged in dialogue with an administrator within the range of once-per-week to once-per-semester, about 72% of participants reported that they collaborated with peers somewhere within the range of once-per-day to once-per-month, and about 79% stated that they engaged in self-reflection within the range of once-per-day to once-per-quarter.

The findings for H2 indicated that most teachers were at least minimally conversing with their principals about grading practices, collaborating with their peers, and reflecting on their personal practices. Also, the findings implied that because there were not significant differences in the types of collaboration in which teachers participated, yet they were not implementing best-grading practices, these supports might not be the cause of their failure to do so. Rather, there could be other unknown variable(s), such as other types of supports, influencing teachers' likelihood to implement best practices.

ANOVA was applied to check for differences in agreement on the topics listed in H3. After analyzing the data, no differences were established topic-to-topic. No single

topic or set of topics stood out as having stronger perceptions than the entire grouping of survey prompts considered. When ANOVA was applied to the percentage of agreement versus the percentage of disagreement for survey prompts 6 through 41 and prompts 43 through 56, question-to-question, in attempt to identify whether strong differences in perception existed, significant differences were also not found. Though trends of strong agreement were found in responses that allowed a grouping of prompts that received similar response patterns from participants, there were no significant differences found in a comparison of the level of response rates when considered topic-to-topic.

For prompts 6 through 41, a *z*-test for difference in proportions was applied to check for differences in agreement versus disagreement on the topics covered by individual questions (see Table 9). Prompts 6, 7, 8, 14, 17, 20, 24, and 37 indicated significant disagreement. Complete disagreement was indicated by participants for prompts 19, 21, and 26. Significant agreement was indicated by participants for the following prompts: 9, 11, 12, 13, 16, 18, 22, 25, 27, 28, 29, 30, 31, 32, 33, 35, 36, and 41. There were no significant differences found for prompts 10, 15, 23, 34, 38, 39, and 40.

Therefore, the data showed that 100% of survey participants completely disagreed with the inclusion of the non-academic factors of a student's gender, physical appearance, and personality when grading (prompts 19, 21, and 26). Similarly, approximately 97% of participants significantly disagreed that they included the non-academic factor of a student's GPA (prompt 17). However, about 94% responded that they agreed with the inclusion of the non-academic factor of students losing points for not completing all parts of an assignment (prompt 16). These findings demonstrated that, in accordance with expert recommendation, all teachers surveyed excluded certain physical characteristics of

students in grades. However, most did not follow expert recommendation of excluding

student responsibility in grades or reporting it separately from academic factors.

A *z*-test for difference in proportions was also applied to check for differences in

agreement versus disagreement on the topics covered by individual questions for prompts

43 through 54 (see Table 10). Significant agreement was indicated by participants for

prompts 43, 47, 49, 50, 52, 53, 55, and 56. There were no significant differences found

for prompts 48 and 54. About 88% of participants indicated that they agreed with prompt

49 (an academic factor) that grades should be a direct reflection of student mastery of

course learning standards. Conversely, 50% of participants responded to prompt 51 (an

academic factor) that they disagreed that their gradebook included a field that indicated

each student's mastery of learning competencies. These findings showed that, while the

majority of teachers agreed with the expert recommendation that grades should reflect

student achievement of standards, half of teachers reported they did not have a gradebook

field by which to indicate such. Therefore, this support was not available for use by all

teachers.

After analyzing the data, H4 was supported because the analysis on responses to

the survey prompts yielded many significant relationships (see Appendix M). For

prompts 6 through 41 and 43 through 56, potential relationships between participant

responses were analyzed with the application of PPMCC. Strong relationships were

found between a number of prompts (see Table 11). Prompt 27 (a non-academic factor)

had the largest number of positive relationships. Responses to prompt 27 (the

consideration of a student's effort when grading) indicated strong agreement with

prompts 12, 13, 18, 22, 25, 28, 29, 30, 32, 33, 35, and 41. When participants agreed that

they considered a student's effort, they also agreed with the following non-academic prompts: 12 (teachers' exercise of personal judgment in grades), 13 (assigning grades for assignments other than assessments), 18 (assigning grades for class participation), 22 (assigning grades for homework), 25 (lowering grades for turning an assignment in after its due date), and 28 (the belief that grades should reward students for high academic achievement). Those that agreed with prompt 27 also agreed with the following academic prompts: 29 (considering student progress towards the course learning standards when grading), 30 (assigning individual grades for students working in groups), 32 (using project-based grading), 33 (assigning grades for project-based assignments), 35 (conferencing with students about their grades throughout each grading period), and 41 (the belief that teachers should be accountable for following best-grading practices). Agreement with this mixed range of prompts showed that teachers had conflicting values about separating academic and non-academic factors in grades.

After analyzing the data for RQ1, there were four supports found to work best for teachers trying to implement best-grading practices. First, was collaboration with peers. Approximately 47% of participants agreed that their districts provided collaboration time for teachers in the same content area, about 59% indicated that their districts did not provide common, collaborative planning time for teachers in the same content area, around 53% of respondents stated their districts did not provide collaborative time for teachers on early-release days, and about 62% stated their districts failed to provide collaboration time for teachers on late-start days. However, about 82% of respondents stated that their districts provided collaboration time for teachers on professional development (PD) days. An implication of these results is that if the majority of teachers

were able to collaborate on PD days, then districts should take action, such as providing

more PD days throughout the school year or institute protected times and/or supports

wherein there would be greater opportunities for teachers to engage in self-reflection and

collaboration in attempt to move towards greater consistency and implementation of best

practices.

Second, collaboration with administrators was found to work as one of the best

supports. About 53% of respondents indicated they engaged in dialogue with their

administrator once-per-quarter to once-per-semester. About 12% of respondents stated

that this never happened, whereas about 15% stated it only happened once-per-year.

Another 21% of respondents stated that they engaged in dialogue with their administrator

once-per-week to once-per-month. Therefore, an implication of these results was that if

administrators and teachers were habitually sharing their observations about grading

practices, districts might be able to develop cultures wherein there was frequent and

meaningful examination of building and district-wide grading practices that promoted

feedback, alignment, buy-in, and accountability. Engagement by both parties increased

the likelihood of long-term sustainability and promoted a collaborative vision of grading

that encouraged active roles for all stakeholders. Provided that all interviewed teachers

communicated a desire for more informal and formal opportunities by which to

collaborate with their leaders about grading practices, it was further implied that any such

supports that districts were to put into place to create these opportunities would be

welcomed by teachers.

Third, pre-service training was found to work as one of the best supports. While

about 35% of survey respondents indicated their teacher education undergraduate

coursework included training on best-grading practices, all interviewed teachers stated

their training left much to be desired. Additionally, approximately 33% of survey

respondents disagreed that they received such training, while about 32% of respondents

neither agreed nor disagreed. The results showed that a gap then-currently existed

between teachers that did receive grading practice training in their preparation programs

and teachers that did not. Despite inconsistent pre-service training, all teachers in this

study possessed the power of autonomy when assigning grades. The results implied,

therefore, that a need then-currently existed to consistently educate pre-service teachers

about best-grading practices that may serve to improve both grading practices and the

credibility of the profession as a whole. As stated by DuFour (2011), no dictionary

existed wherein a professional was defined as "someone who can do whatever he or she

pleases" (p. 58).

Fourth, PD was found to work as one of the best supports for teachers trying to

implement best-grading practices. Less than one-half (approximately 44%) of survey

respondents indicated their districts provided PD about best-grading practices. About

38% stated that their districts did not. While about 82% of respondents stated they were

able to collaborate with their peers on days dedicated to PD, the data showed that most

districts were not providing their teachers with PD about best-grading practices. These

results indicated that there was a then-current need for districts to implement training

about best-grading practices in attempt to maximize the utilization of collaboration time.

It was therefore implied that district-provided PD may help correct the lack of pre-service

training teachers received about best-grading practices by disseminating information

about such.

After analyzing the data for RQ2, there were 13 academic factors found that teachers used when grading. The factors consisted of the following: Student progress towards achieving the course learning goals, project-based assignments with individual grading for such, individual projects, individual grades for group work, summative assessments, and assignments that have been re-done, self-reflection via comparison with colleagues in the same content area about grading practices, conferences with students about their grades throughout each grading period, that teachers should be accountable for following best-grading practices, that grades should be a direct reflection of student master of the course learning standards, that grades are a means by which to provide feedback to stakeholders about student progress of the course learning standards, and that grades are an indication of how students should perform on state exams. The three academic factors that teachers reported using the most were (a) grades should be a direct reflection of student master of the course learning standards; (b) individual projects and; and (c) holding conferences with students about their grades throughout each grading period. Findings from the surveys and interviews showed that while most teachers agreed that the three aforementioned factors should be emphasized when calculating grades, teachers had varying opinions about the inclusion of other academic factors. The findings also showed that teachers utilized varying combinations of such factors. Consistent with the findings of previous research, these findings implied that teachers' then-current grading practices could produce grades resulting in considerable variations from teacher-to-teacher. It was further implied that although the desire to follow best practice existed, teachers had not been properly educated about which factors should be reflected in grades and which should be reported separately.

After analyzing the data for RQ3, it was found that teachers chose to use academic factors when grading for four reasons pertaining to their beliefs. Firstly, they believed that grades should include several measurable academic factors that served to indicate a student's overall academic aptitude. Secondly, teachers believed grades should be a direct reflection of student mastery of the course learning standards. Thirdly, they believed grades should provide feedback about student progress of the course learning standards to stakeholders. Fourthly, teachers believed grades served as an indication of how students should be expected to perform on state exams. It was, therefore, implied that teachers were aware that academic factors were important to the act of grading; but, it was not because they were formally educated about the role of academic factors in grades. Instead, teachers learned to emphasize the role of academic factors when grading, because that was how they were graded by their own teachers when they were students, and/or they developed their own practices based on classroom experience.

After analyzing the data for RQ4, it was found that teachers chose to use 10 non-academic factors when grading. They were as follows: Student effort, student ability level, extra credit, zeroes for missing assignments, formative assessments, assignment completion, class participation, homework, the act of using grades as a reward for high academic achievement and; student responsibility. Additionally, some survey participants further noted that they included attendance, a student's personal level of interest in the subject, behavior, and a student's home environment when grading. Therefore, consistent with the findings of previous research, the findings of this study implied that teachers valued non-academic factors and chose to combine them with academic factors in their grading calculations. However, the inclusion of particular non-academic factors and the

extent to which they were emphasized varied depending on the individual teacher.

After analyzing the data for RQ5, it was found that teachers chose to include non-academic factors when grading for reasons pertaining to their individual beliefs surrounding the importance of using grades, as a means by which to indicate a student's sense of responsibility and their perceived ability levels. Consistent with the findings of previous research, it was therefore implied that teachers' beliefs about these topics and the act of including non-academic factors in grades to express them further complicated the validity and meaning of grades.

Personal Reflections

Before I started this study, I tried to think of topics to investigate that would help provide greater perspective about teachers' experiences in their classrooms. As a teacher myself, I know how meaningful it is to self-reflect about one's practices and how meaningful grades can be to all stakeholders. Teachers' sentiments about what it really means to teach (including both the celebrations and the tribulations associated with such) are something to be valued. Often times, these thoughts are communicated teacher-to-teacher in the faculty lounge. It was my hope that I could embark on a study that honored teachers by providing insight about a topic that consistently weighs on their minds - grading practices. Therefore, I wanted this study to serve as a resource to inform districts and administrators about the then-current state of teachers' grading practices in order to help bridge the gap between expert recommendation and existing practice and improve the system of grading for the benefit of all stakeholders.

As stated by Webster (2011), "We ask teachers to complete the critical task of grading without much preparation or guidance" (pp. 192-193). All of the teachers I

interviewed shared that they had received little to no formal training in their pre-service

teacher education programs. Instead, they had all individually crafted their practices

relying mostly on their personal experiences. Teacher #3 best summarized the

interviewed teachers' sentiments when she stated,

> I know what the experts say best-grading practices are but I don't feel like they
>
> really get how widespread the problem of lackadaisical students has become. So I
>
> use their recommendations to inform me, but I don't think of them as laws. If
>
> more practicing teachers were involved in the creation of these recommendations,
>
> I think more teachers would be willing to implement them.

In the interviews and on the survey, teachers candidly expressed their experiences

with grading and their recommendations as to how the practice might be improved,

sharing what they felt were the positive and negative aspects of grading. They openly

detailed the factors they included in their grading practices and the reasons why they

considered such factors important. It was apparent in both forums how personal, and

ultimately emotional, the topic of grading practices was to teachers. It was also apparent

how morally invested they were in the process, due to their awareness of the potentially

life-long consequences grades can have for students. For all of these reasons, I think that

the teachers who participated in the survey and interviews served as agents of change by

engaging in this investigation of grading practices.

The findings of this study connected with the review of literature. The review of

literature revealed teachers' historical struggle with the ethical dilemmas pertaining to the

practice of grading and their consideration of multiple factors when calculating grades.

Findings, such as teachers' desire to exercise their personal judgment when grading and

an overall lack of formal training about best-grading practices were also consistent with the review of literature. However, perhaps the most surprising finding was how great the lack of pre-service training about best-grading practices truly was for teachers. On a similar note, a perspective not found in the literature was how little time and types of support about grading (such as dialogue between teachers and administrators, PD, and collaborative time) were provided by districts. A considerable amount of teachers articulated that they desired improvements in these areas and felt that their practices would benefit from such. The act of changing pre-service training programs and the amount of time and types of support provided by districts are necessary steps to take to initiate the process of implementation of best-grading practices with fidelity.

Recommendations to the Program

As a result of this study, I recommend that administrators at each participant school district begin to reflect on the topic of grading practices by taking some time to consider how personal and emotional teachers' grading practices are and how deeply important they are to them. I hope that administrators are able to understand these feelings in order to better comprehend the struggle that teachers undergo on a daily basis in trying to grade their students' coursework in a manner that is judicial, accurate, and consistent. As evidenced by the participants' responses and opinions in this study, teachers were astutely aware of the importance of grades and the factors included when calculating them. However, there was considerable variation in the factors teachers included when calculating their grades. All interview participants indicated that they were aware of such inconsistencies in teachers' grading practices, whether it was among teachers in the same content area, building, or district. Despite the immediate and

potential life-long consequences grades can carry for their students, teachers did not receive adequate training to accurately complete the task.

As a result, teachers found themselves at the center of an ethical dilemma presented by the act of grading that was as constant as it was taxing. This dilemma with grading was also something that cannot be solved solely by teachers. However, by taking specific actions, administrators, school districts, and collegiate teacher preparation programs are able to work together to help alleviate this burden and correct the gap that exists between expert-recommended grading practices and those currently being used by teachers.

Therefore, after contemplating the emotional and ethical demands teachers experience when grading, I recommend that administrators review student grades and course syllabi of all of the teachers in their buildings. Districts should also review any handbooks and formal, written grading policies they might have. Reviewing such documents allows administrators to gain a better understanding of the values their teachers possess that translate into the factors they choose to include in their grades. As encouraged by Noll (2008), educational professionals should explore problems to determine "what is it and how it got that way" (p. xvii). Doing so provides administrators with perspective regarding the source and state of issues surrounding grading practices in their schools and better prepares them to combat problems.

I further recommend that administrators engage in frequent dialogue with teachers about their grading practices. Doing so facilitates the establishment of a positive environment wherein administrators are able to provide teachers with guidance, opportunities for input, and constructive feedback. It also encourages teacher buy-in and

self-reflection about grading practices. Administrators could conduct interviews or focus groups with teachers to obtain their opinions about current practices as well as their aspirations and suggestions pertaining to implementation of best practices.

Districts should strive to put supports in place that provide administrators with the knowledge to accurately assess teachers' grading practices. This could be accomplished by providing principals with PD about best-grading practices, conducting book studies, or outlining literature reviews and theories of action. Districts should also provide additional guidance to teachers. Such supports include reserved time for teachers to collaborate with peers and administrators about grading practices, as well as providing PD about best-grading practices.

While these might seem like relatively simple solutions, the literature and findings of this study revealed that the issues surrounding teachers' grading practices are anything but simple. Rather, the subject of grading practices is enmeshed in other complex educational issues, such as "What is the purpose of school? What do we believe about children? What do we believe about learning?" (Webster, 2011, p. 193). Teachers should regularly exercise self-reflection about their grading practices as a good faith effort in working toward the establishment of uniformly accepted answers to these questions with the goal of widespread implementation of best-grading practices. Sustainable grading reform is more likely to occur after teachers have actively scrutinized traditional practices, thereby having divorced themselves from the historical practice wherein academic and non-academic factors are combined into a singular, composite A through F grade.

The majority of teachers in this study were experienced veterans. Their willingness to discuss their experiences with grading throughout their years of teaching, as well as the factors they included when calculating grades revealed that they are self-aware of the importance of their roles in the grading process. Their willingness to discuss obstacles associated with grading and their suggestions for improvement indicated that they desire systematic change. This desire, coupled with the literature, reveals an overall need for leadership that requires administrators, teachers, and districts to collaborate with one another and establish types of support in order to successfully alter the practice of grading.

While teachers in this study communicated that they are willing to self-reflect on their grading practices and seek opportunities to learn about best-grading practices, the process of change is complicated by the fact that grading practices are so embedded in each teacher's history (both as a student, as well as the leader of a class). Their grading practices were also embedded within their school's culture (Schein, 1990). Therefore, the shift from then-current practice to sustainable reform demanded that administrators and teacher leaders possess an understanding of adaptive change and demonstrated a commitment to spearhead such change (Webster, 2011). According to Heifetz and Linsky (2002), adaptive change addressed problems that required people in an institution letting go of "things they hold dear: daily habits, loyalties, ways of thinking" (p. 2) for the potential reward of securing a better outcome.

Consequently, in order for grading system change to be effective, it is imperative that administrators assume a personal level of responsibility for the problem and a commitment to fixing it. Administrators must answer the call for change and navigate

their teachers to a system of better grading practices by taking the initiative to introduce change and create the incentive to change. Heifetz and Linsky (2002) maintained that leaders should focus teachers' attention on the tough issues, forcing them to assume ownership for "tackling and solving those issues, and by bringing conflicts occurring behind closed doors out into the open" (p. 6). Teachers in this study also communicated the need for administrators to collaborate with other administrators in their districts to ensure consistency in grading practice expectations from grade-to-grade and building-to-building. Therefore, I recommend that administrators share their observations and experiences regarding their teachers' grading practices with one another to facilitate a collective vision of implementation of best practices.

Within this collective vision, it is necessary that collegiate institutions also assume an active role in the review and administrative oversight of the implementation of best-grading practices. This can be accomplished by colleges and teacher preparation programs revamping their curriculum to incorporate coursework about best-grading practices and the concerns pertaining to traditional practices. Colleges should work with schools to coordinate in-service, PD, and undergraduate and graduate course options that expand on the tenets of best-grading practices. Schools and practicing teachers and administrators could provide colleges and teacher preparation programs with additional information regarding grading issues, concerns, and recommendations. Developing a collaborative relationship between institutions increases the overall level of knowledge about grading for all stakeholders and lends itself to the establishment of yet another level of support for leaders and practitioners attempting to improve the grading system.

Recommendations for Future Research

Findings from the survey, interviews, and review of literature revealed that a gap existed between expert-recommended grading practices and then-current practice. Teachers in this study articulated a need to learn about grading practices in their pre-service training programs and a need for district-provided time and support in which to collaborate and implement best-grading practices. In addition, they also discussed obstacles such as inconsistent expectations about grading practices from administrators. All participants emphasized a willingness to serve as agents of change and a desire to see more practicing teachers involved in discussions about change. Teachers' involvement would likely lead to greater implementation of best practices. This was evidenced by the fact that participants spoke about their desire to find a balance between their feeling of uncertainty, while still maintaining their autonomy.

Santoro (2011) maintained that while a considerable amount of research was conducted to explore the individual characteristics that influenced teachers' abilities to withstand the challenges of the profession, few studies explored how the quality of the work - the practice of teaching - affected teachers. Santoro (2011) suggested that an exploration of the demoralization of teachers could provide a valuable perspective from which to analyze the challenges then-presently clouding the profession. Therefore, in attempt to further explore how best to structure the educational environment and identify which types of support structures facilitate the implementation of best-grading practices, it might be beneficial for future research to explore how the actual practice of teaching affected teachers. Such research might contribute to the field by expanding upon Santoro's (2011) argument that it was important to understand that "the moral dimension

of teaching is not only about cultivating individual teachers' dispositions towards good work but also about structuring the work to enable practitioners to do good within its domain" (p. 3). And, that "Good work not only serves society, it yields personal gratification and provides vital sustenance to the profession itself" (Santoro, 2011, p. 4). In other words, the more the educational community understands the types of support needed for teachers to implement best-grading practices with fidelity, the greater the benefit to student learning and society as a whole.

It might also be beneficial for future research to analyze existing curriculum in the courses of undergraduate and graduate level teacher education programs, as it pertains to best-grading practices to determine how to better train teachers. Since interview participants indicated they had little to no grading-practice training, an exploration of reform strategies in this area could better prepare teachers to implement best-grading practices. Colleges and school districts might also consider forming a steering committee or focus group of teachers, administrators, and teacher education students to review expert-recommended grading practices with the goal of establishing a working model to aide other teachers, administrators, and student teachers.

Similarly, future studies might consider whether administrators received grading practice training in their specialist programs. Educating administrators about best-grading practices helped to navigate what Reeves (2008) called "the last frontier of individual teacher discretion" (p. 86) - grading. An exploration of both teacher and administrator preparatory programs lends itself to establishing an educational environment wherein teachers feel comfortable dialoging with administrators and other leaders about grading practices. Such an environment could enable teachers to assume a more active role in

school policy-making about grading practices; thereby, helping to revert the current trend

of teacher demoralization and increase the implementation of best-grading practices. This

type of environment fostered the identification of the consequences of inconsistent

grading practices, which Reeves (2008) stated helped to create "a sense of urgency;"

thereby, inciting educational leaders to rise to the occasion and accept the challenge of

reforming grading practices (p. 87).

Guskey and Bailey (2010) found that administrators sought to improve the

consistency of teachers' grading practices, but they were concerned that introducing

changes in this area would threaten teachers' sense of autonomy. Given that teachers

placed such importance on their ability to exercise personal judgment when grading, it

might be necessary for districts to assess teachers' grading practices as part of their

summative or student teaching evaluations. Future studies might want to consider this

recommendation and examine whether doing so might facilitate collaboration among

peers and dialogue between teachers and administrators about best-grading practices to

build support for reform strategies.

This study focused on the grading practices of grade 7 through 12 core content

area teachers in four southeast Missouri schools. Additional research is necessary to build

a more comprehensive understanding of teachers' grading practices in K-6 schools and in

grade 7 through 12 non-core content areas. More research in such areas would increase

understanding of grading. Schools seeking to implement grading reform strategies could

use data collected in this study as baseline data. However, while the data collected in this

study provided insight into then-current grading practices, a larger sample size in future

studies might increase perspective about grading, as well as the generalizability of

findings. Examining teacher demographics beyond those included in this study, such as race and language, would allow researchers to gain a more diverse perspective in understanding grading.

Conclusion

Given how personal teachers' grading practices are, and how deeply they are embedded within a school's culture, it is of value to understand who teachers are - their demographics and their values - in order to better understand how they feel about current grading practices. Kotter and Cohen (2002) stated that people change "because they are shown a truth that influences their feelings" (p. 1). In order to achieve sustainable reform, administrators and districts need to lead teachers and show them that change is necessary to ensure judicial, accurate, and consistent grading. This must be done in a way that connects to the hearts of teachers, drawing on their core value of doing what is best for students. This study helps to present the compelling need for change by providing data concerning teachers' demographics and insight regarding their views on current practices.

The over-arching theme of the research was to provide information regarding the factors that teachers consider when grading and the best types of support that districts could provide to facilitate teachers' implementation of best-grading practices. The five research questions focused on teacher demographics, the specific academic and non-academic factors teachers included when grading, and the reasons why they included such factors. All of these things must be considered by teacher preparatory programs in order to better tailor their curriculum to produce teachers that have been formally trained in best-grading practices. Principals and district leaders should also take these things into consideration when reflecting on how to begin bridging the gap that existed between

grading practice and best practice, at the time of this study. Professional development

coordinators must too consider these things in order to provide teachers with the proper

supports by which to reinforce the implementation of best practices.

Finally, above all else, this study serves to offer teacher testimony regarding the

hope for change in grading practices and the commitment to bettering the educational

experience for students.

References

Airasian, P. W. (1997). *Classroom assessment* (3rd ed.). New York, NY: McGraw-Hill.

Airasian, P. W. (2005). *Classroom assessment: Concepts and applications* (5th ed.). NewYork, NY: McGraw-Hill Higher Education.

Aliaga, M., & Gunderson, B. (2002). *Interactive statistics*. Upper Saddle River, NJ: Prentice Hall.

Allen, J. (2005). Grades as valid measures of academic achievement of classroom learning. *The Clearing House, 78*(5), 218-223. doi:10.3200/TCHS.78.5.218-223

Anderman, E. M., Maehr, M. L., & Midgley, C. (1999). Declining motivation after the transition to middle school: Schools can make a difference. *Journal of Research and Development in Education, 32*(3), 131-147. doi:10.2307/1170696

Ashbaugh, E. J., & Chapman, H. B. (1925). Report cards in American cities. *Educational Research Bulletin, 4*(14), 289-293.

Austin, S., & McCann, R. (1992, April). *"Here's another arbitrary grade in your collection": A statewide study of grading policies.* Paper presented at the annual meeting of the American Education Research Association, San Francisco, CA.

Bailey, M. (2012). *The relationship between secondary school teacher perceptions of grading practices and secondary school teacher perceptions of student motivation.* Retrieved from http://pqdtopen.proquest.com/doc/1011481355. html?FMT=AI

Bailey, J., & McTighe, J. (1996). Reporting achievement at the secondary level: What and how. In T. R. Guskey (Ed.), *Communicating student learning: ASCD*

yearbook 1996 (pp. 119-140). Alexandria, VA: Association for Supervision and Curriculum Development.

Barnes, S. (1985). A study of classroom pupil evaluation: The missing link in teacher education. *Journal of Teacher Education, 36*(4), 46-49.

Barth, R. (2003)*. Lessons learned*. Thousand Oaks, CA: Corwin Press.

Biddle, B. J., & Anderson, D. S. (1986). *Methods, knowledge, and research on teaching.* In M. C. Wittrock (Ed.), Handbook of research on teaching (3rd ed., pp. 230-252). New York, NY: Macmillan.

Bloom, B. S. (1968). Learning for mastery. *Evaluation comment, 1*(2), 5-10.

Bluman, A. G. (2008). *Elementary statistics: A brief version* (4th ed.). Boston, MA: McGraw-Hill/Higher Education.

Bogdan, R. C., & Biklen, S. K. (2007). *Qualitative research for education: An introduction to theory and methods* (5th ed.) New York, NY: Pearson.

Brookhart, S. (1991). Grading practices and validity. *Educational Measurement: Issues and Practice 10*(1), 35-36. doi:10.1111/j.1745-3992.1991.tb00182.x

Brookhart, S. (1993). Teachers' grading practices: Meaning and values. *Journal of Educational Measurement*, *30*(2), 123-142. doi:10.1111/j.1745-3984.1993.tb01070.x

Brookhart, S. (1994). Teachers' grading: Practice and theory. *Applied Measurement in education, 7*(4), 279-301. doi:10.1207/s15324818ame0704_2

Brookhart, S. (1999). The Art and Science of Classroom Assessment: The Missing Part of Pedagogy. *ASHE-ERIC Higher Education Report 27*(1). Washington, DC: The

George Washington University Graduate School of Education and Human Development.

Brookhart, S. (2004). *Grading*. Upper Saddle River, NJ: Pearson, Merrill, Prentice-Hall.

Brookhart, S. (2009). *Grading* (2nd ed.). Upper Saddle River, NJ: Pearson, Merrill, Prentice-Hall.

Brookhart, S. (2011). Starting the conversation about grading. *Educational Leadership, 69*(3), 10-14.

Brookhart, S., & Nitko, A. (2008). *Assessment and grading in classrooms*. New York, NY: Prentice Hall.

Bruckman, K. (2010). *Influence of problem behavior and teacher tolerance on student grades.* Retrieved from http://drum.lib.umd.edu/bitstream/handle/1903/10278/ Bruckman_umd_0117N_11107.pdf?sequence=1&isAllowed=y

Burke, R. J. (1968) Student reactions to course grade. *The Journal of Experimental Education*, *36*(4), 11–13.

Bursuck, W., Polloway, E., Plante, L., Epstein, M., Jayanthi, M., & McConeghy, J. (1996). Report card grading and adaptations: A national survey of classroom practices. *Exceptional Children, 62*(4)*,* 1-10.

Buzzelli, C. A., & Johnston, B. (2002). *The moral dimensions of teaching: Language, power, and culture in classroom interaction.* New York, NY: Routledge/Falmer Press.

Canady, R. L., & P. R. Hotchkiss. (1989). It's a good score! Just a bad grade. *Phi Delta Kappan, 71*(1), 68–71. Retrieved from http://www.jstor.org/stable/ 20404061

Carlson, L. (2003). Beyond assessment to best-grading practice: Practical guidelines. In J. Wall & G. Walz (Eds.) *Measuring up: Assessment issues for teachers, counselors, and administrators* (Ch. 36). ERIC Clearinghouse version. ED480379. Retrieved from https://files.eric.ed.gov/fulltext/ED480379.pdf

Chansky, N. M. (1964). A note on the grade point average in research. *Educational and Psychological Measurement*, *24*(1), 95-99.

Chiekem, E. (2015). Grading practice as valid measures of academic achievement of secondary school students for national development. *Journal of Education and Practice, 6*(26), 24-29.

Chastain, K. (1990). Characteristics of graded and ungraded compositions. *Modern Language Journal, 74*(1)*,* 10–14.

Cizek, G. (2003). *Detecting and preventing classroom cheating: Promoting integrity in assessment*. Thousand Oaks, CA; Sage Publications.

Cizek, G. J., Fitzgerald, S. M., & Ranchor, R. E. (1996). Teachers' assessment practices: Preparation, isolation, and the kitchen sink. *Educational Assessment, 3*(2), 159-179. doi:10.1207/s15326977ea0302_3

Clymer, J. B., & Wiliam, D. (2007). Improving the way we grade science. *Educational Leadership, 64*(4), 36-42.

Colvin, S. S. (1912). Marks and the marking system as an incentive to study. In J. A. Laska & T. Juarez (Eds.), *Grading and marking in American schools* (pp. 85-92)*.* Springfield, IL: Charles C. Thomas.

Conley, D. (2000, April). *Who is proficient: The relationship between proficiency scores and grades.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Costello, C., & McKillop, B. (2000). Dealing with lates and absences. *Orbit*, *30*(4), 43-46.

Creswell, J. W. (1998). *Qualitative inquiry and research design: Choosing among five traditions.* Thousand Oaks, CA: Sage.

Creswell, J. (2002). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research.* Upper Saddle River, NJ: Merrill Prentice Hall.

Creswell, J. W. (2009) *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Thousand Oaks, CA: Sage.

Creswell, J. W., & Clark, Plano Clark, V. L. (2007). *Designing and conducting mixed methods research.* Thousand Oaks, CA: Sage.

Cross, L. H., & Frary, R. B. (1996, April). *Hodgepodge grading: Endorsed by students and teachers alike.* Paper presented at the annual meeting of the National Council on Measurement in Education, New York, NY.

Cross, L. H., & Frary, R. B. (1999). Hodgepodge grading: Endorsed by students and teachers alike. *Applied Measurement in Education*, *12*(1), 53-72. doi:10.1207/s15324818ame1201_4

Davis, J. (2015). Give teachers time to collaborate. *Education Week, 35*(4), 26-27.

Dee, T. S. (2006). The why chromosome. *Education Next*, *6*(4). 68-75.

Dockery, E. R. (1995). Better grading practices. *Education Digest*, *60*(5), 34-37.

Duckworth, A., & Seligman, M. (2006). Self-discipline gives girls the edge: Gender in

    self-discipline, grades, and achievement test scores. *Journal of Educational*

    Psychology, 96(1), 198-208.

Dueck, M. (2014). *Grading smarter not harder: Assessment strategies that motivate kids*

    and help them learn. Alexandria, VA: Association for Supervision and

    Curriculum Development.

Dufour, R. (2011). Work together: But only if you want to. *Phi Delta Kappan, 92*(5), 57-

    61.

DuFour, R., DuFour, R., Eaker, R., & Many, T. (2006) *Learning by doing: A handbook*

    for professional learning communities at work. Bloomington, IN: Solution Tree.

Dunn, J. L., & Allen, J. (2009, Winter). Holding schools accountable for the growth of

    nonproficient students: Coordinating measurement and accountability.

    Educational Measurement: Issues, and Practice, 28(4), 27-41. doi:10.1111/

    j.1745-3992.2009.00160.x

Dyd, D. (2012). *Teacher perceptions of middle level grading practices.* [Doctoral

    Dissertation]. Madison, WI: Edgewood College).

Ebel, R. L. (1979). *Essentials of educational measurement* (3rd ed.). Englewood Cliffs,

    NJ: Prentice Hall.

Ebel, R. A., & Frisbie, D. A. (1986). *Essentials of educational measurement* (4th ed.).

    Englewood Cliffs, NJ: Prentice-Hall.

Edwards, C. (1999). Let's end the grading game. *The Clearing House*, *72(5),* 260-263.

Effective grading practices in the middle school and high school environments. (2011).

    Hanover Research. Retrieved from http://www.wawm.k12.wi.us/UserFiles/

Servers/Server_1212649/File/C%20&%20I/Grading%20Articles/Grading%20Pra

ctices%20Hanover%20Research%20Summary.pdf

Eison, J. A., & Pollio, H. R., & Milton, O. (1986). *Making sense of college grades: Why*

the grading system does not work and what can be done about it. San Francisco,

CA: Jossey-Bass.

Erickson, J. (2010). Grading practices: The third rail. *Principal Leadership*, *10*(7), 22-24.

Erickson, J. (2011). A call to action transforming grading practices. *Principal*

Leadership, *12*(1), 42-46.

Feldmesser, R. A. (1971). *The positive functions of grades.* Paper presented at the annual

meeting of the American Educational Research Association, New York, NY.

Feistritzer, E. C. (2011). Profile of teachers in the United States. *National Center for*

Education Information. Retrieved from https://www.edweek.org/media/pot2011

final-blog.pdf

Fleming, N. (1999). *Assessment matters in higher education: Choosing and using diverse*

approaches. London, UK: McGraw-Hill Education.

Frary, R., Cross, L., & Weber, L. (1992, April). *Testing and grading practices and*

opinions in the nineties: 1890s or 1990s? Paper presented at the annual meeting

of the National Council on Measurement in Education, San Francisco, CA.

Frary, R. B., Cross, L. H., & Weber, L. J. (1993). Testing and grading practices and

opinions of secondary teachers of academic subjects: Implications for instruction

in measurement. *Educational Measurement: Issues and Practice, 12*(3), 23-30.

Friedman, S. J., & Manley, M. (1991, April). *Grading practices in the secondary school: Perceptions of the stakeholders.* Paper presented at the annual meeting of the National Council on Measurement in Education, Chicago, IL.

Gottfredson, G. D. (1981). Schooling and delinquency. In S. E. Martin, L. B. Sechrest, & R. Redner (Eds.), *New directions in the rehabilitation of criminal offenders* (pp. 424-469). Washington, DC: National Academy Press.

Gottfried, A. E. (1985). Academic intrinsic motivation in elementary and junior high school students. *Journal of Educational Psychology, 77*(6), 631-645. doi:10.1037/0022-0663.77.6.631

Green, K. H., & Emerson, A. (2007). A new framework for grading. *Assessment & evaluation in Higher education*, *32*(4), 495-511.

Greene, B. A., Miller, R. B., Crowson, H. M., Duke, B. L., & Akey, K. L. (2005). Predicting high school students' cognitive engagement and achievement: Contributions of classroom perceptions and motivation. Contemporary Educational Psychology, *29*(4), 462-482.

Greville, E. (2009). A rose by any other name: Grading and assessment. *Assessment update: Progress, trends and practices in higher education, 21*(5), 1-2, 13.

Griswold, P. A. (1993). Beliefs and inferences about grading elicited from student performance sketches. *Educational Assessment*, *1*(4), 311-328.

Guskey, T. (1996). *Communicating student learning: The ASCD yearbook.* Alexandria, VA: Association for Supervision and Curriculum Development.

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

Guskey, T. (2002). Computerized gradebooks and the myth of objectivity. *Phi Delta Kappan, 83*(10), 775-780.

Guskey, T. (2004). 0 alternatives. *Principal Leadership*, 5(2), 49-53.

Guskey, T. (2006). Making high school grades meaningful. *Phi Delta Kappan, 87(9)*, 670-675.

Guskey, T. (2008). *Practical solutions for serious problems in standards-based grading.* Thousand Oaks, CA: Corwin Press.

Guskey, T. (2009). *Bound by tradition: Teachers' views of crucial grading and reporting issues.* Paper presented at the Annual Meeting of the American Educational Research Association. San Francisco, CA.

Guskey, T. (2011). Grading reform. *Educational Leadership*, *69*(3), 17-21.

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

Guskey, T., & Bailey, J. (2001). *Developing grading and reporting systems for student learning.* Thousand Oaks, CA: Corwin Press.

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

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

Hardre', P. L., Chen, C. H., Huang, S. H., Chiang, C. T., Jen, F. L., & Warden, L. (2006). Factors affecting high school students' academic motivation in Taiwan. *Asia Pacific Journal of Education, 26(2),* 189-207.

Hargreaves, A., & Fullan, M. (2013). The power of professional capital. *JSD-The Learning Professional, 34(3),* 36-39.

Hawkins, J. (2010). *The impact of teacher perception of student effort and student attitude on grading of secondary students.* Retrieved from http://escholarshare. drake.edu/bitstream/handle/2092/1447/dd2010JPH.pdf?sequnce=1

Heifetz, R. A., Grashow, A., & Linsky, M. (2009). *The practice of adaptive leadership: Tools and tactics for changing your organization and the world.* Cambridge, MA: Harvard Business Press.

Heifetz, R. A., & Linsky, M. (2002). *Leadership on the line*. Cambridge, MA: Harvard Business School Press.

Heflebower, J., Hoegh, K., & Warrick, P. (2014). *A school leader's guide to standards-based grading*. Bloomington, IN: Solution Tree Press.

Hendrickson, J., & Gable, R. (1997). Collaborative assessment of students with diverse needs: Equitable, accountable, and effective grading. *Preventing School Failure*, *41*(4), 159-164.

Hills, J. R. (1991). Apathy concerning grading and testing. *Phi Delta Kappan, 72(7)*, 540-545.

Hinshaw, S. P. (1992). Externalizing behavior problems and academic underachievement in childhood adolescence: Causal relationships and underlying mechanisms. *Psychological Bulletin*, *111*(1), 127-155.

Holmes, L., & Smith, L. (2003). Student evaluations of faculty grading methods. *Journal of Education for Business*, 78(6), 318-323.

Hord, S. M. (1997). *Professional learning communities: Communities of continuous inquiry and improvement.* Austin, TX: Southwest Educational Development Laboratory.

Hord, S., & Sommers, W. (2008). *Leading professional learning communities: Voices from research and practice.* Thousand Oaks, CA: Corwin Press.

Hyde, G. M. (1897). Should "marks" be withheld? *Journal of Education, 45*(6), 92-93.

Ingersoll, R. (2003). *Who controls teachers' work? Power and accountability in America's schools.* Cambridge, MA: Harvard University Press.

Ingersoll, R., & Merrill, L. (2010). Who's teaching our children? *Educational Leadership*, *67*(8), 14-20.

Johnson, D. W., & Johnson, R. T. (2002). *Meaningful assessment: A manageable and cooperative process.* Boston, MA: Allyn and Bacon.

Johnson, W., McGue, M., & Iacono, W. G. (2005). Genetic and environmental influences on academic achievement trajectories during adolescence. *Developmental Psychology, 42*(3), 514-532.

Jongsma, K. S. (1991). Rethinking grading practices. *The Reading Teacher, 45*(A), 318-320.

Jung, L., & Guskey, T. (2010). Grading exceptional learners. *Educational Leadership, 67*(5), 31-35.

Kain, D. L. (1996). Looking beneath the surface: Teacher collaboration through the lens of grading practices. *Teachers College Record, 97*(Summer), 569-587.

Kohn, A. (1994). Grading: The issue is not how but why. *Educational Leadership, 52(2)*, 38-41.

Kohn, A. (1999). *Punished by rewards: The trouble with gold stars, incentive plans, A's, praise, and other bribes.* Boston, MA: Houghton Mifflin.

Kotter, J. P., & Cohen, D. S. (2002). *The heart of change: Real-life stories of how people change their organizations.* Boston, MA: Harvard Business School Press.

Lekholm, A., & Cliffordson, C. (2008). Discrepancies between school grades and test scores at individual and school level: Effects of gender and family background. *Educational Research and Evaluation, 14*(2), 181-199.

Leana, C. R. (2011, Fall). The missing link in school reform. *Stanford Social Innovation Review*, *9*(4), 34.

Leyrer, B. (2015). *The evolution of grading experiences, policies, and practices of high school teachers.* Retrieved from http://ezproxy.lindenwood.edu:2048/login?url= https://search-proquest-com.ezproxy.lindenwood.edu/docview/1699340335? accountid=12104

Linn, R. L., Baker, E. L., & Betebenner, D. W. (2002). Accountability systems: Implications of requirements of the No Child Left Behind Act of 2001. *Educational Researcher, 31*(6), 3-16. doi:10.3102/0013189X031006003

Little, J. W. (1982). Norms of collegiality and experimentation: Workplace conditions of school success. *American Educational Research Journal, 19*(3), 325–340.

Lomax, R. (1996). On becoming assessment literate: An initial look at pre-service teachers' beliefs and practices. *The Teacher Educator, 31*(4), 292-303.

Lunenburg, F. (2010). Creating a professional learning community. *National Forum of Educational Administration and Supervision Journal, 27*(4), 1-7.

Malouff, J. M. (2008). Bias in grading. *College Teaching, 56*(3), 191-192.

Malouff, J. M., & Thorsteinsson, E. B. (2016). Bias in grading: A meta-analysis of

    experimental research findings. *Australian Journal of Education, 60*(3), 245–256.

    doi:10.1177/0004944116664618

Malouff, J. M., Stein, S. J., Bothma, L. N., Coulter, K., & Emmerton, A. J. (2014).

    Preventing halo bias in grading the work of university students. *Cogent*

    Psychology, 1(1). Retrieved from http://cogentoa.tandfonline.com/doi/full/10.

    1080/23311908.2014.988937

Manke, M. P., & Loyd, B. H. (1990, April). *An investigation of non-achievement related*

    factors influencing teachers' grading practices. Paper presented at the annual

    meeting of the National Council on Measurement in Education, Boston.

Many, T. (2009, May/June). Make time for collaboration. *Texas Elementary Principals*

    and Supervisors Association News, 2009(May/June), 8-9.

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

    for Supervision and Curriculum Development.

Marzano, R. J. (2006). *Classroom assessment and grading that work*. Alexandria, VA:

    Association for Supervision and Curriculum Development.

Marzano, R. J. (2007). *The art and science of teaching*. Alexandria, VA: Association for

    Supervision and Curriculum Development.

Marzano, R. J. (2010). *On excellence in teaching* (10th ed.). Bloomington, IN: Solution

    Tree Press.

McMillan, J. H. (2001). Secondary teachers' classroom assessment and grading practices.

    Educational Measurement, *20*(1), 20-32.

McMillan, J. H. (2002). Elementary school teachers' classroom assessment and grading practices. *Educational Measurement: Issues and Practice, 20(1)*, 20-32.

McMillan, J. H. (2007). Classroom assessment: *Principles and practice for effective* (4th ed.). Boston, MA: Pearson Education.

McMillan, J. H. (2008). *Assessment essentials for standards-based education* (2nd ed.). Thousand Oaks, CA: Sage Publications.

McMillan, J. H. (2012). *SAGE handbook of research on classroom assessment.* Thousand Oaks, CA: Sage Publications.

McMillan, J. H., & Lawson, S. (2001). *Secondary teachers' classroom assessment and grading practices.* Richmond, VA: Virginia Commonwealth University and The Metropolitan Educational Consortium.

McMillan, J. H., & Schumacher, S. (2006). *Research in education: A conceptual introduction.* New York, NY: Longman.

McMillan, J. H., Myran, S., & Workman, D. (1999, April). *The impact of mandated state-wide testing on teachers' classroom assessment and instructional practices.* Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada. Retrieved from http://eric.ed.gov/ERICDocs/data/ ericdocs2sql/content_storage_01/0000 019b/80/17/98/49.pdf

McMillan, J. H., Myran, S., & Workman, D. (2002). Elementary teachers' classroom assessment and grading practices. *The Journal of Educational Research, 95*(4), 203-213.

McMunn, N., Schenck, P., & McColskey, W. (2003, April). *Standards-based assessment, grading, and reporting in classroom: Can district training and support change*

teacher practice? Paper presented at the Annual Meeting of the American

Educational Research Association, Chicago, IL.

McTighe, J., & Ferrara, S. (1998). *Assessing learning in the classroom*. Student

Assessment Series. Washington, DC: National Educational Association

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: A sourcebook for

new methods* (2nd ed.). Thousand Oaks, CA: Sage.

Moll, M. (1998). A brief history of grading. *Teacher News Magazine, 11*(3), 1-2.

Montgomery, D. (2009, November 19). Half-dozen districts sue over Texas law

prohibiting minimum grades. *Dallas Star Telegram.* Retrieved from http://www.

star-telegram.com

Muijs, D. (2010). *Doing quantitative research in education using SPSS.* Thousand Oaks,

CA: Sage.

Mulford, B. (2003). *School leaders: Changing roles and impact on teacher and school

effectiveness.* Paper presented to the Education and Training Policy Division,

OECD, for the Activity Attracting, Developing and Retraining Effective

Teachers, Adelaide, SA.

National Center for Education Statistics. (2019). Fast facts. Retrieved from

https://nces.ed.gov/fastfacts/display.asp?id=28

National Staff Development Council. (2009). *NSDC's definition of professional

development*. Retrieved from http://www.nsdc.org/standfor/definition.cfm

National Women's History Museum. (2007). T*he history of women and education.*

Retrieved from http://www.nwhm.org/exhibits/education/Introduction.html

Nitko, A. J. (2001). *Educational assessment of students* (3rd ed.). Upper Saddle River, NJ: Merrill/Prentice Hall.

Nodding, N. (2008). *When school reform goes wrong*. New York, NY: Teachers College Press.

Noll, J. (2008). *Taking sides: Clashing views on educational issues.* New York, NY: McGraw-Hill.

O'Connor, K. (2002). *How to grade for learning: Linking grades to standards* (2nd ed.). Glenview, IL: Skylight.

O'Connor, K. (2007). *A repair kit for grading: 15 fixes for broken grades*. Princeton, NJ: Educational Testing Service.

O'Connor, K. (2009). *How to grade for learning: Linking grades to* standards (3rd ed.). Thousand Oaks, CA: Corwin Press.

O'Connor, K. (2012). *How to grade for learning by using 15 fixes for broken grades.* [PowerPoint presentation]. Retrieved from http://www.gfps.k12.mt.us/sites/ default/ files/Ken O'Connor GFPS Aug 13, 2012.pdf

O'Connor, K., & Wormeli, R. (2011). Reporting student learning. *Educational Leadership,* 69(3), 40-44.

Ornstein, A. C. (1994). Grading practices and policies: An overview and some suggestions. *NASSP Bulletin, 78*(559), 55–64.

Panagos, K. (2011). *Professional learning communities: A phenomenological study of year-one implementation.* [Doctoral dissertation]. St. Louis, MO: Saint Louis University.

Patrick, M. D. (2009). *Addressing the lack of male elementary school teachers: Factors that influence grade-level preference*. Retrieved from http://digitalcommons. liberty.edu/cgi/viewcontent.cgi?article=1276&context=doctoral

Patton, M. Q. (2008). *Utilization-focused evaluation* (4th ed.). Thousand Oaks, CA: Sage.

Polloway, E., Epstein, M., Bursuck, W., Roderique, T., McConeghy, J., & Jayanthi, M. (1994). Classroom grading: A national survey of policies. *Remedial and Special Education, 15*(3), 162-170.

Popham, W. (2008). *Transformative assessment.* Alexandria, VA: Association for Supervision and Curriculum Development.

Randall, J., & Engelhard, G. (2009). Differences between teachers' grading practices in elementary and middle schools. *The Journal of Educational Research, 102*(3), 175-185.

Randall, J., & Engelhard, G. (2010). Examining the grading practices of teachers. *Teaching and Teacher Education, 26*(7), 1372-1380.

Raywid, M. A. (1993). Finding time for collaboration. *Educational leadership*, *51*(1), 30-34.

Reeves, D. (2006). *A funny thing happened on the way to the learning leader: How to focus school improvement for better results.* Alexandria, VA: Association for Supervision and Curriculum Development.

Reeves, D. (2008). Effective grading practices. *Educational Leadership*, *65*(5), 85-87.

Reeves, D. (2010). *Standards, assessment, and accountability: Real questions from educators with real answers.* Englewood, CA: The Leadership and Learning Center.

Reeves, D. (2011). Taking the grading conversation public. *Educational Leadership*, *69*(3), 76-79.

Richmond, E. (2008, February 4). A floor for failing grades: Parents, educators debate whether kids should get 50 points for doing nothing. *Las Vegas Sun*. Retrieved from https://lasvegassun.com/news/2008/feb/04/floor-failing-grades/

Rinsland, H. D. (1937). *Constructing tests and grading in elementary and high school subjects.* Oxford, UK: Prentice-Hall.

Rom, M. C., & Musgrave, P. (2014). Political outcome bias in grading: Identifying problems and suggesting solutions. *Journal of Political Science Education*, *10*(2), 136-154.

Santoro, D. A. (2011). Good teaching in difficult times: Demoralization in the pursuit of good work. *American Journal of Education, 118*(1), 1-23. doi:10.1086/662010

Santoro, D. A. (2012, May). Teacher demoralization and teacher burnout: Why the distinction matters [AJE Forum]. Retrieved from http://www.ajeforum.com/teacher-demoralization-and-teacher-burnout-why-the-distinction-matters/

Scantlebury, K. (2009, December 23). Gender bias in teaching. In E. Anderman (Ed.). *Psychology of classroom learning: An encyclopedia* (221-224). Detroit, MI: Macmillan Reference USA.

Schein, E. H. (1990, February). Organizational culture. *American Psychologist, 45*(2), 109-119. doi:10.1037//0003-066X.45.2.109

Schneider, J., & Hutt, E. (2013). Making the grade: A history of the A-F marking scheme. *Journal of Curriculum Studies*, *46*(2), 1-24.

Scriffiny, P. (2008). Seven reasons for standards-based grading. *Educational Leadership, Expecting Excellence*, *66*(2), 70-74.

Shanahan, K. B. (2011). *Elementary teachers' grading practices: Does the reality reflect the rubric?* Retrieved from http://drum.lib.umd.edu/bitstream/handle/1903/12276/ Shanahan_umd_0117E_12720.pdf;sequence=1

Simon, S. B., & Bellanca, J. A. (1976). *Degrading the grades myths: Primer of alternatives to grades and marks.* Washington, DC: Association for Supervision and Curriculum Development.

Sinclair, R. L., & Ghory, W. I. (1987). *Reaching marginal students: A primary concern for school renewal.* Berkeley, CA: McCutchan.

Slavin, R. E. (2007). *Educational research in an age of accountability.* Boston, MA: Pearson.

Snyder, T. D. (1993). 120 years of American education: A statistical portrait. National Center for Educational Statistics. Washington, DC: U.S. Government Printing Office.

Southeast Missouri State University. (2019). About RPDC [Regional Professional Development Center]. Retrieved from https://semo.edu/rpdc/about.html

Stacy, M. (2013). Teacher-led professional development: Empowering teachers as self advocates. *The Georgia Social Studies Journal, 3*(1), 40-49. Retrieved from https: //coe.uga.edu/assets/files/misc/gssj/Stacy-2013.pdf

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

Starch, D., & Elliott, E. C. (1913a). Reliability of grading work in history. *School Review, 21*(10)*,* 254-295.

Starch, D., & Elliott, E. C. (1913b). Reliability of grading work in mathematics. *School Review, 21*(4), 676-681.

Stiggins, R. (2007). Assessment through the student's eyes. *Educational Leadership,64*(8), 22–26.

Stiggins, R. J. (1993). Teacher training in assessment: Overcoming the neglect. In S. L. Wise (Ed.), *Teacher training in measurement and assessment skills* (pp. 27-40). Lincoln, NE: Buros Institute of Mental Measurements.

Stiggins, R. J. (1999). Are you assessment literate? *The High School Journal, 6*(5), 20-23.

Stiggins, R. J. (2001). Student-involved classroom assessment. (3rd ed.). Upper Saddle River, NJ: Merrill/Prentice Hall.

Stiggins, R. J. (2002). Assessment crisis: The absence of assessment for learning. *Phi Delta Kappan, 83*(10), 758-765.

Stiggins, R. J. (2005). *Student-involved assessment for learning.* Upper Saddle River, NJ: Pearson/Prentice Hall.

Stiggins, R. J. (2008). *Assessment for learning, the achievement gap, and truly effective schools.* Presented at the Educational Testing Service and College Board Conference, Educational Testing in America: State Assessments, Achievement Gaps, National Policy and Innovations, Washington, DC.

Stiggins, R. J., & Conklin, N. F. (1992). *In teachers' hands.* Nova Iorque, BR: State University of New York Press.

Stiggins, R. J., & Duke, D. L. (1991). District grading policies and their potential impact

      on at-risk students. Paper presented at the annual meeting of the American

      Educational Research Association, Chicago, IL.

Stiggins, R. J., Frisbie, D. A., & Griswold, P.A. (1989). Inside high school grading

      practices: Building a research agenda. *Educational Measurement: Issues and*

      Practice, 8(2), 5-14.

Strein, W., & Meshbesher, N. (2006). Grades and grading. In G. Bear and K. Minke

      (Eds.). *Children's needs: Psychological perspectives III* (Ch. 45). Bethesda, MD:

      National Association of School Psychologists.

Tetreault, D. (2005). Administrative technology: New rules, new tools – A pilot study of

      Excelsior software's electronic gradebook solution reveals the impact and time-

      saving qualities of administrative software. *Technological Horizons in Education*

      Journal, *32*(9), 39-44.

Tomlinson, C. A. (2001). *How to differentiate instruction in mixed-ability classroom.*

      Alexandria, VA: Association for Supervision and Curriculum Development.

Tomlinson, C. A. (2003). *Fulfilling the promise of the differentiated classroom.*

      Alexandria, VA: Association for Supervision and Curriculum Development.

Tomlinson, C. A. (2005, Summer). Grading and differentiation: Paradox or good

      practice. *Theory Into Practice, 44*(3), 262-269. doi:10.1207/s15430421tip

      4403_11

Tomlinson, C. A. (2007). Learning to love assessment. *Educational Leadership*, *65*(4),

      8-13.

Tomlinson, C. A., & McTighe, J. (2006). *Integrating differentiated instruction and understanding by design.* Alexandria, VA: Association for Curriculum and Supervision Development.

Tomlinson, C. A., & Moon, T. R. (2013). *Assessment and student success in a differentiated classroom*. Alexandria, VA: Association for Supervision and Curriculum Development.

Truog, A. L., & Friedman, S. J. (1996). *Evaluating high school teachers' written grading policies form a measurement perspective.* Paper presented at the annual meeting of the National Council on Measurement in Education, New York, NY. Retrieved from www.naesp.org/resources/2/Leadership.../ LC2006v4n2a3.pdf

Tsang, K. K., & Liu, D. (2016). Teacher demoralization, disempowerment, and school administration. *Qualitative Research in Education, 5*(2), 200-225.

Wang, D. (2013). *The demoralization of teachers: Crisis in a rural school in China.* Lanham, MD: Lexington Books.

Webster, K. (2011). *High school grading practices: Teacher leaders' reflections, insights, and recommendations.* Retrieved from http://pqdtopen.proquest.com/ doc/929134496.html?FMT=AI

Wenglinski, H. (2000). *How teaching matters: Bringing the classroom back into discussions of teacher quality.* Princeton, NJ: Educational Testing Service.

Wiggins, G. (1994). Toward better report cards. *Educational Leadership, 52*(2), 28-37. Retrieved from http://www.ascd.org/publications/educational-leadership/oct94/ vol52/num02/Toward-Better-Report-Cards.aspx

Wiggins, G. (1996). Honesty and fairness: Toward better grading and reporting. In T. R.

      Guskey (Ed.), *Communicating student learning: 1996 Yearbook of the association*

      for supervision and curriculum development (pp. 141-76). Alexandria, VA:

      Association for Supervision and Curriculum Development.

Wiggins, G., & McTighe, J. (2006). Examining the teaching life. *Educational*

      Leadership, 63(6), 26-29.

Wiggins, G., & McTighe, J. (2007). *Schooling by design: Mission, action, and*

      achievement. Alexandria, VA: Association for Supervision and Curriculum

      Development.

Wilson, M. (2004). *Towards coherence between classroom assessment and*

      accountability. Chicago, IL: The University of Chicago Press.

Winger, T. (2005). Grading to communicate. *Educational Leadership, 63*(3), 61-65.

Winger, T. (2009). Grading what matters. *Educational Leadership, 67*(3), 73-75.

Wormeli, R. (2006). Accountability: Teaching through assessment and feedback, not

      grading. *American Secondary Education*, *34*(3), 14-27.

Yesbeck, D. (2011). *Grading practices: Teachers' considerations of academic and non-*

      academic factors. Retrieved from http://scholarscompass.vcu.edu/cgi/ view

      content.cgi?article=3585&context=etd

Zoeckler, L. (2007). Moral aspects of grading: A study of high school English teachers'

      perception. *American Secondary Education*, *35*(2), 83-102.

Appendix

Appendix A: Recruitment Statement for Survey Research Participation

Research Sites: Washington, Adams, Jefferson, and Madison School Districts.
1. You are invited to participate in a research study conducted by Jenelle Lee, a
2. graduate student at Lindenwood University, under the guidance of Dr. Bob Steffes.

3. The title of this study is *An Investigation of Patterns Between Demographic Characteristics of Teachers and Grading Practices.* The purpose of this study is to investigate patterns between teacher demographic characteristics, grading, and the amount of time and support systems provided by a district. Specific demographics that will be explored include: gender, age, content area, years of service and educational level. Specific support systems to be explored include: conversation with administration, collaboration with peers, professional development and teacher education program training. Findings from this study
4. will contribute to existing research in the educational community, specifically in the areas of identification of factors used in grading practices by teachers possessing particular demographic characteristics and the amount of time and support systems provided by districts to facilitate implementation of best-grading practices.

5. Your participation in this study will include

    a. Participants will be asked to complete a survey that will keep all responses anonymous. Surveys are able to be completed via Qualtrics at the participants' convenience. Surveys should take no longer than 20 minutes to complete.

    b. The expected duration of the study is 12 months. The researcher hopes to have the study completed by October 2017. The duration of each subject's participation will be no more than 9 months. The researcher will only contact participants asking for survey and interview privileges.

6. There are minimal risks to you as a participant in the study. Misuse of time is considered a possible risk, as it would take away from valuable instructional time in the classroom. The researcher acknowledges your time restraints and professional commitments and has taken steps to maximize data collection efforts as well as schedule data collection during specified non-instructional times. By organizing and establishing timelines both the researcher and the participants will be able to benefit from time commitments and obligations that are necessary for this study.

7. The results of this study may be published in scientific research journals or presented at professional conferences. However, your name and identity will not

be revealed and your record will remain confidential. To maintain confidentiality, the researcher will code all interview and observational data, with a master list kept secured and separate. All data that is collected will be stored in a locked file cabinet to protect the privacy of participants. Once the study has been completed all data collection material, responses, and records will be stored in a locked file cabinet for three years.

8. Participation in this study will benefit you by providing valuable information regarding factors used in grading practices by teachers possessing particular demographic and administrative leadership at grade 7-12 SE RPDC member schools in Missouri. Participants in this study can use the data obtained from this study to improve teacher grading practices initiatives and assessment of student academic achievement, planning for professional development, as well as maximize teacher collaboration and student learning. This investigation will enable the researcher to expand upon existing research and make contributions to the educational community, specifically in the areas of identification of factors used in grading practices by teachers possessing particular demographic characteristics and the amount of time and support systems provided by districts to facilitate implementation of best-grading practices.

9. You can choose not to participate. If you decide not to participate, there will not be a penalty to you or loss of any benefits to which you are otherwise entitled. You may withdraw from this study at any time.

10. If you have any questions about this research study, you can contact Jenelle Lee at 417-850-2270 or the supervising faculty member, Dr. Bob Steffes, at 636-494-7400. If you have any questions about your rights as a research participant, you can call the Lindenwood University Institutional Review Board at (636) 949-2000.

I have read this consent form and have been given the opportunity to ask questions. I will also be given a copy of this consent form for my records. I consent to my participation in the research described above.

_____   _____
Participant's Signature   _____Date   Participant's Printed Name

_____   _____
Signature of Researcher   _____Date   Researcher Printed Name

Appendix B: Recruitment Statement for Telephone Interview Research

Participation

Research Sites: Washington, Adams, Jefferson, and Madison School Districts.

1. You are invited to participate in a research study conducted by Jenelle Lee, a graduate student at Lindenwood University, under the guidance of Dr. Bob Steffes.

2. The title of this study is *An Investigation of Patterns Between Demographic Characteristics of Teachers and Grading Practices.* The purpose of this study is to investigate patterns between teacher demographic characteristics, grading, and the amount of time and support systems provided by a district. Specific demographics that will be explored include: gender, age, content area, years of service and educational level. Specific support systems to be explored include: conversation with administration, collaboration with peers, professional development and teacher education program training. Findings from this study will contribute to existing research in the educational community, specifically in the areas of identification of factors used in grading practices by teachers possessing particular demographic characteristics and the amount of time and support systems provided by districts to facilitate implementation of best-grading practices.

3. Your participation in this study will include

   c. Participants will be asked to complete an interview that will keep all responses anonymous. Interviews will be completed via telephone at the participants' convenience. Interviews should not exceed 30 minutes. Participants who consent to participating in an interview will be contacted by the researcher by telephone at an agreed upon time.

   d. The expected duration of the study is 12 months. The researcher hopes to have the study completed by October 2017. The duration of each subject's participation will be no more than 9 months. The researcher will only contact participants asking for survey and interview privileges.

4. There are minimal risks to you as a participant in the study. Misuse of time is considered a possible risk, as it would take away from valuable instructional time in the classroom. The researcher acknowledges your time restraints and professional commitments and has taken steps to maximize data collection efforts as well as schedule data collection during specified non-instructional times. By organizing and establishing timelines both the researcher and the participants will be able to benefit from time commitments and obligations that are necessary for this study.

5. The results of this study may be published in scientific research journals or presented at professional conferences. However, your name and identity will not be revealed and your record will remain confidential. To maintain confidentiality, the researcher will code all interview and observational data, with a master list kept secured and separate. All data that is collected will be stored in a locked file cabinet to protect the privacy of participants. Once the study has been completed all data collection material, responses, and records will be stored in a locked file cabinet for three years.

6. Participation in this study will benefit you by providing valuable information regarding factors used in grading practices by teachers possessing particular demographic and administrative leadership at grade 7-12 SE RPDC member schools in Missouri. Participants in this study can use the data obtained from this study to improve teacher grading practices initiatives and assessment of student academic achievement, planning for professional development, as well as maximize teacher collaboration and student learning. This investigation will enable the researcher to expand upon existing research and make contributions to the educational community, specifically in the areas of identification of factors used in grading practices by teachers possessing particular demographic characteristics and the amount of time and support systems provided by districts to facilitate implementation of best-grading practices.

7. You can choose not to participate. If you decide not to participate, there will not be a penalty to you or loss of any benefits to which you are otherwise entitled. You may withdraw from this study at any time.

8. If you have any questions about this research study, you can contact Jenelle Lee at 417-850-2270 or the supervising faculty member, Dr. Bob Steffes, at 636-494-7400. If you have any questions about your rights as a research participant, you can call the Lindenwood University Institutional Review Board at (636) 949-2000.

I have read this consent form and have been given the opportunity to ask questions. I will also be given a copy of this consent form for my records. I consent to my participation in the research described above.

_____   _____

Participant's Signature   _____Date   Participant's Printed Name

_____   _____

Signature of Researcher   _____Date   Researcher Printed Name

Appendix C: Request for Permission to Survey

District Grade 7-12 Core Content Teachers

Dear Superintendent,

I am a graduate student at Lindenwood University, and I am inviting grade 7-12 core content area teachers in member schools of the Southeast Regional Professional Development Center (SE RPDC) to participate in a research study. The purpose of this study is to investigate patterns between teacher demographic characteristics, grading, and the amount of time and support systems provided by a district. Specific demographics that will be explored include: gender, age, content area, years of service and educational level. Specific support systems to be explored include: conversation with administration, collaboration with peers, professional development and teacher education program training. Findings from this study will contribute to existing research in the educational community, specifically in the areas of identification of factors used in grading practices by teachers possessing particular demographic characteristics and the amount of time and support systems provided by districts to facilitate implementation of best-grading practices.

I will be using Qualtrics Survey Generator to disperse the survey and obtain data. Once the teachers submit their survey through Qualtrics, I will ensure that participants' names and identities will not be revealed and their records will remain confidential. Data will be collected for the purpose of identifying patterns in the types of factors used in grading practices by teachers of particular demographic orientations located in 3 counties that are members of the SE RPDC.

I respectfully request that you please forward the email to the grade 7-12 core content area teachers in your school district. Thank you in advance for your assistance regarding this research.

Please find attached hereto my Recruitment Statement for Research Participation and a copy of my survey questions.

Kind regards,

Jenelle Lee

Appendix D: Survey

Please select one answer to each of the following demographic characteristics.

D.C. 1 Years of teaching experience: 0-5   5-10 10-15 15+
D.C. 2 Gender: M   F
D.C. 3 Highest degree earned: BS   Master Specialist Doctorate
D.C. 4 Content area: ELA S.S. Science Math
D.C. 5 Age: 20-29 30-39 40-49 50+

Please rate the following statements on a scale of 1 to 5, where 1=Strongly Disagree and 5=Strongly Agree.

6. I consider a student's behavior when grading
7. Grades are a means to punish students for low academic achievement
8. I assign points for bringing daily supplies to class
9. I consider a student's ability level when grading
10. I assign points just for completing an assignment
11. I allow students to earn extra credit
12. Teachers should be able to exercise a degree of personal judgment in assigning grades
13. I assign grades for assignments other than assessments
14. I consider a student's involvement in extracurricular activities when grading
15. I assign students zeroes for missing assignments
16. Students lose points if they do not complete all parts of an assignment
17. I consider a student's GPA when grading
18. I assign grades for class participation
19. I consider a student's gender when grading
20. I have lowered a student's grade for demonstrating poor behavior
21. I consider the physical appearance of a student when grading
22. I assign grades for homework
23. I assign grades for taking notes
24. I consider the aesthetic appearance of the students' work when grading
25. I assign students a lower grade for turning in an assignment after its due date
26. I consider a student's personality when grading
27. I consider a student's effort when grading
28. Grades should reward students for high academic achievement
29. I consider student progress towards achieving the course learning standards when grading
30. I assign individual grades for students working in groups
31. I assign grades for individual projects
32. I use project-based grading in my classroom
33. I assign grades for project-based assignments
34. I include student portfolios in my grading
35. I conference with students about their grades throughout each grading period
36. I require students to set personal learning goals for the course
37. I require my students to reflect on their achievement of the course learning standards in a journal or notebook

38. My district provides time for teachers in the same content area to meet to align grading practices

39. My district provides professional development about best-grading practices

40. I received training on best-grading practices in my undergraduate coursework

41. Teachers should be accountable for following best-grading practices

42. How often does your administrator engage in dialogue with you about best-grading practices?

Never   1x daily   1x weekly   1x monthly   1x quarterly   1x per semester   1x yearly

43. I consider how my grading practices compare to other teachers' grading practices in the same content area as myself

44. My district provides common planning time for teachers in the same content area to promote collaboration

45. My district provides collaboration time for teachers on early release days

46. My district provides collaboration time for teachers on late start days

47. My district provides collaboration time for teachers on professional development days

48. I regularly collaborate with grade level teachers during common plan time

49. I believe grades should be a direct reflection of student mastery of learning standards

50. My grades provide feedback about student progress of the learning standards to stakeholders

51. My gradebook includes a field that indicates each student's mastery of learning competencies

52. I allow students to re-do assignments

53. I allow students to re-take a test if they fail it

54. I grade summative assessments, not formative assessments

55. I believe that my grades are an indication of how students should perform on state exams

56. I weigh summative assessments heavier than formative assessments when grading

57. What other factors do you consider when grading a student's work?

58. How often do you collaborate with your peers? (Please select one response.)

Never   1x daily   1x weekly   1x monthly   1x quarterly   1x per semester   1x yearly

59. What topics do you discuss with your peers if you collaborate with them?

60. How often do you self-reflect about the accuracy and fairness of your grading practices? (Please select one response.)

Never   1x daily   1x weekly   1x monthly   1x quarterly   1x per semester   1x yearly

Appendix E: Coding for Teacher Survey

Table E1

Coding for Teacher Survey

Survey Question Numbers	Code
29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 43, 44, 45, 46, 47, 47, 49, 50, 51, 52, 53, 54, 55, 56	Academic factors
6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28	Non-Academic factors
1, 2, 3, 4, 5	Demographics
42, 57, 58, 59, 60	Other

Appendix F: Request for Permission to Interview

District Grade 7-12 Core Content Teachers

Dear Superintendent,

I am a graduate student at Lindenwood University, and I am inviting grade 7-12 core content area teachers in member schools of the Southeast Regional Professional Development Center (SE RPDC) to participate in a research study. The purpose of this study is to investigate patterns between teacher demographic characteristics, grading, and the amount of time and support systems provided by a district. Specific demographics that will be explored include: gender, age, content area, years of service and educational level. Specific support systems to be explored include: conversation with administration, collaboration with peers, professional development and teacher education program training. Findings from this study will contribute to existing research in the educational community, specifically in the areas of identification of factors used in grading practices by teachers possessing particular demographic characteristics and the amount of time and support systems provided by districts to facilitate implementation of best-grading practices.

I will conduct interviews via telephone at the interviewee's convenience. Interviews will not exceed 30 minutes.

I respectfully request that you please forward the interview questions to the grade 7-12 core content area teachers in your school district that you believe would be ideal candidates to interview. Then, please email me contact information for the teachers willing to participate. I will email them a consent form to sign and arrange a time to conduct the interview. Thank you in advance for your assistance regarding this research.

Please find attached hereto my Recruitment Statement for Research Participation and a copy of my interview questions.

Kind regards,

Jenelle Lee

Appendix G: Interview Questions

1. Tell me about the academic factors that you use in your grading practices. Why have you chosen to incorporate these factors?

2. Tell me about the non-academic factors that you use in your grading practices. Why have you chosen to incorporate these factors?

3. Tell me about any other factors you might consider when grading. Why have you chosen to incorporate these factors?

4. Tell me about your grading practice training.

5. Do you feel that the implementation of measurement experts' recommended grading practices is effective at fairly and accurately measuring student academic achievement, please explain your answer?

6. How would you describe the support of the implementation of expert recommended grading practices from the teachers?

7. How would you describe the support of the implementation of expert recommended grading practices from administrators?

8. Tell me about the dialogue between you and your administrator about grading practices.

9. Tell me about the self-reflection you undergo regarding your grading practices.

10. How do your grading practices compare to your peers?

11. Do you feel that you have adequate time and resources to implement expert recommended grading practices, please explain your answer?

12. What changes, if any, would you suggest that could make teachers more effective at improving grading practices?

13. Tell me about the professional development your district has provided regarding grading practices.

14. Tell me about the opportunities your district provides that is specifically intended for teacher collaboration.

Appendix H: Thank You Letter

Dear Educational Professional,

Thank you for your participation in this study. I would like to express my sincere appreciation for your time and efforts. Please know that findings from this study will contribute to existing research in the educational community, specifically in the areas of teacher grading practices and the time and supports districts provide to facilitate teacher implementation of best-grading practices. I look forward to analyzing the data I collected from your district as a result of your participation.

Again, thank you and the best of luck with future endeavors.

Kind regards,

Jenelle Lee

Appendix I: Certificate from NIH IRB Training for Student

Certificate of Completion

The National Institutes of Health (NIH) Office of Extramural Research certifies that **Jenelle Lee** successfully completed the NIH Web-based training course "Protecting Human Research Participants".

Date of completion: 09/05/2015.

Certification Number: 1838678.

Appendix J: Initial Contact Letter

Dear Superintendent,

My name is Jenelle Lee and I am a doctoral student at Lindenwood University. I am also a teacher at Scott City High School. I am reaching out to area school districts belonging to the Southeast Regional Professional Development Center (SE RPDC) in hopes that they will participate in my study.

At this stage of my dissertation, I am only asking that you provide your district's intent to participate in my study. Once I gain IRB approval, I will be sending you additional information such as the link to my online survey to forward to your teachers. The purpose of my study is to investigate patterns between grade 7-12 core content area teacher demographic characteristics, grading, and the amount of time and support systems provided by a district. After participants complete my survey, I will be conducting interviews via telephone at the interviewee's convenience. Interviews will not exceed 30 minutes. If you would like, I can email you the results of the study after its conclusion. Please find my survey questions attached hereto.

Per Lindenwood University's requirement, please copy and paste the following statement on your district letterhead and sign and date it to let me know if you accept or decline my invitation. Please send it back to me by Friday, January 13th via fax or email. Fax: 573-264-2608, Attn: Jenelle Lee. If you have any questions, feel free to contact me by email or on my cell phone at 417-850-2270. Thank you for your consideration.

Kind regards,

Jenelle Lee, doctoral student Lindenwood University and teacher at Scott City High School

I _____, Superintendent of _____ School District hereby give / do not give (please indicate one or the other) my approval for Jenelle Lee to conduct her doctoral study.

_____   _____
Signature   Date

Appendix K: IRB Approval

Date: Mar 3, 2017

To: Jenelle Lee

From: Lindenwood University Instructional Review Board

Project Title: [1006288-1] An Investigation of Patterns Between Demographic
Characteristics of Teachers and Grading Practices

Principal Investigator: Jenelle Lee

Submission Type: New Project
Date Submitted: February 13, 2017

Action: APPROVED
Effective Date: March 3, 2017
Review Type: Expedited Review

Appendix L: ANOVA for Prompts 6-56

Table L1

ANOVA SUMMARY: Prompt 6-56

Groups	Count	Sum	Average	Variance
# 6	5	1	0.2	0.022237
# 7	5	0.9999	0.19998	0.122573
# 8	5	0.9999	0.19998	0.081066
# 9	5	1	0.2	0.023966
# 10	5	1.0001	0.20002	0.006223
# 11	5	1	0.2	0.03175
# 12	5	0.9999	0.19998	0.035643
# 13	5	0.9999	0.19998	0.048616
# 14	5	0.9999	0.19998	0.101388
# 15	5	1	0.2	0.004932
# 16	5	1	0.2	0.096197
# 17	5	1	0.2	0.107006
# 18	5	1	0.2	0.021804
# 19	5	1	0.2	0.2
# 20	5	1	0.2	0.064632
# 21	5	1	0.2	0.2
# 22	5	1	0.2	0.038679
# 23	5	1	0.2	0.010552
# 24	5	0.9999	0.19998	0.06377
# 25	5	0.9999	0.19998	0.013582
# 26	5	1	0.2	0.172329
# 27	5	0.9999	0.19998	0.02051
# 28	5	0.9999	0.19998	0.032618
# 29	5	1	0.2	0.030019
# 30	5	1	0.2	0.046017
# 31	5	1	0.2	0.07501
# 32	5	1	0.2	0.014886
# 33	5	0.9999	0.19998	0.026565
# 34	5	0.9999	0.19998	0.020067
# 35	5	0.9999	0.19998	0.042995
# 36	5	0.9999	0.19998	0.028726
# 37	5	0.9999	0.19998	0.010987
# 38	5	1	0.2	0.004068
# 39	5	1.0001	0.20002	0.002334
# 40	5	1.0001	0.20002	0.004927

# 41	5	1	0.2	0.028292
#43	5	0.9999	0.19998	0.026565
#44	5	1	0.2	0.020509
#45	5	1	0.2	0.03175
#46	5	1	0.2	0.04818
#47	5	0.9999	0.19998	0.044725
#48	5	0.9999	0.19998	0.017905
#49	5	0.9999	0.19998	0.060289
#50	5	1	0.2	0.03175
#51	5	1	0.2	0.018346
#52	5	0.9999	0.19998	0.020498
#53	5	1.0001	0.20002	0.001902
#54	5	1	0.2	0.004931
#55	5	0.9999	0.19998	0.045589
#56	5	0.9999	0.19998	0.017041

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	3.84E-08	49	7.84E-10	1.75E-08	1	1.418051
Within Groups	8.979788	200	0.044899			
Total	8.979788	249				

Table L2

ANOVA: Single
Factor

ANOVA SUMMARY:
Prompt 6-41

Groups	Count	Sum	Average	Variance
# 6	5	1.0000	0.2000	0.0222
# 7	5	0.9999	0.2000	0.1226
# 8	5	0.9999	0.2000	0.0811
# 9	5	1.0000	0.2000	0.0240
# 10	5	1.0001	0.2000	0.0062
# 11	5	1.0000	0.2000	0.0317
# 12	5	0.9999	0.2000	0.0356
# 13	5	0.9999	0.2000	0.0486
# 14	5	0.9999	0.2000	0.1014
# 15	5	1.0000	0.2000	0.0049
# 16	5	1.0000	0.2000	0.0962
# 17	5	1.0000	0.2000	0.1070
# 18	5	1.0000	0.2000	0.0218
# 19	5	1.0000	0.2000	0.2000
# 20	5	1.0000	0.2000	0.0646
# 21	5	1.0000	0.2000	0.2000
# 22	5	1.0000	0.2000	0.0387
# 23	5	1.0000	0.2000	0.0106
# 24	5	0.9999	0.2000	0.0638
# 25	5	0.9999	0.2000	0.0136
# 26	5	1.0000	0.2000	0.1723
# 27	5	0.9999	0.2000	0.0205
# 28	5	0.9999	0.2000	0.0326
# 29	5	1.0000	0.2000	0.0300
# 30	5	1.0000	0.2000	0.0460
# 31	5	1.0000	0.2000	0.0750
# 32	5	1.0000	0.2000	0.0149
# 33	5	0.9999	0.2000	0.0266
# 34	5	0.9999	0.2000	0.0201
# 35	5	0.9999	0.2000	0.0430
# 36	5	0.9999	0.2000	0.0287
# 37	5	0.9999	0.2000	0.0110
# 38	5	1.0000	0.2000	0.0041

# 39	5	1.0001	0.2000	0.0023
# 40	5	1.0001	0.2000	0.0049
# 41	5	1.0000	0.2000	0.0283

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	2.73E-08	35	7.79E-10	1.51E-08	1	1.5050
Within Groups	7.4198	144	0.0515			
Total	7.4198	179				

Table L3

ANOVA SUMMARY: Prompt 43- 56

Groups	Count	Sum	Average	Variance
#43	5	0.9999	0.2000	0.0266
#44	5	1.0000	0.2000	0.0205
#45	5	1.0000	0.2000	0.0317
#46	5	1.0000	0.2000	0.0482
#47	5	0.9999	0.2000	0.0447
#48	5	0.9999	0.2000	0.0179
#49	5	0.9999	0.2000	0.0603
#50	5	1.0000	0.2000	0.0317
#51	5	1.0000	0.2000	0.0183
#52	5	0.9999	0.2000	0.0205
#53	5	1.0001	0.2000	0.0019
#54	5	1.0000	0.2000	0.0049
#55	5	0.9999	0.2000	0.0456
#56	5	0.9999	0.2000	0.0170

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	1.09E-08	13	8.35E-10	3E-08	1	1.8993
Within Groups	1.5599	56	0.0279			
Total	1.5599	69				

Appendix M: Strong Positive Relationships

Table M1

Q	Strong Positive Relationship (> 0.750) With
# 6	7, 8, 14, 17, 20, 21, 24, 26
# 7	6, 8, 14, 17, 19, 20, 21, 24, 26
# 8	14, 17, 19, 20, 21, 24, 26
# 9	10, 11, 12, 30, 32
# 10	11, 12, 23, 30, 36
# 11	9, 10, 12, 18, 23, 25, 29, 30, 32, 39
# 12	9, 10, 11, 18, 25, 27, 29, 30, 32, 39
# 13	16, 22, 27, 28, 29, 31, 33, 35, 41
# 14	6, 7, 8, 19, 20, 21, 24, 26
# 15	35
# 16	13, 22, 28, 31, 33, 35, 41
# 17	6, 7, 8, 14, 19, 20, 21, 24, 26
# 18	11, 12, 22, 25, 27, 28, 29, 30, 32, 33, 41
# 19	6, 7, 8, 14, 17, 20, 21, 24, 26,
# 20	6, 7, 8, 14, 17, 19, 21, 24, 26
# 21	6, 7, 8, 14, 17, 19, 20, 24, 26
# 22	13, 16, 18, 25, 27, 28, 29, 30, 32, 33, 35
# 23	10, 11, 36, 39
# 24	6, 7, 8, 14, 17, 19, 20, 21, 26
# 25	11, 12, 18, 22, 27, 29, 30, 32
# 26	6, 7, 8, 14, 17, 19, 20, 21
# 27	12, 13, 18, 22,25, 28, 29, 30, 32, 33, 35, 41
# 28	14, 16, 18, 22, 27, 29, 31, 33, 35, 41
# 29	11, 12, 13, 18, 22, 25, 28, 30, 32, 33, 35
# 30	9, 10, 11, 12, 18, 222, 25, 27, 29, 32, 39
# 31	13, 16, 28, 33, 35, 41
# 32	9, 11, 12, 18, 22, 25, 28, 29, 31, 33
# 33	13, 16, 18,22, 27, 28, 29, 32, 35, 41
# 34	31
# 35	13, 15, 16, 27, 28, 29, 31, 33, 41
# 36	10, 23
# 37	34
# 38	
# 39	11, 12, 23, 30
# 40	
# 41	13, 16, 18, 27, 28, 29, 31, 33, 35

42
43   49, 55
44   45, 48
45   44, 46, 51
46   45, 51
47   50, 52
48   44, 51
49   43, 50, 52
50   47, 49, 52
51   45, 46, 48
52   47, 49, 50
53   56
54
55   43
56   53

Appendix N: Strong Inverse Relationships

Table N1

Strong Inverse Relationships

	# 6	# 7	# 8	# 9	# 10	# 11	# 12	# 13
# 6	1							
# 7	0.935395	1						
# 8	0.910889	0.991094	1					
# 9	-0.07642	-0.20413	-0.10881	1				
# 10	-0.35276	-0.38508	-0.27148	0.775281	1			
# 11	-0.1642	-0.12193	0.003457	0.884378	0.811932	1		
# 12	-0.28573	-0.28538	-0.16589	0.923468	0.853443	0.982359	1	
# 13	-0.59331	-0.43477	-0.41064	0.125333	-0.05215	0.30179	0.32646	1
# 14	0.951094	0.998689	0.989576	-0.17182	-0.37179	-0.11121	-0.27062	-0.45341
# 15	-0.83039	-0.90793	-0.91096	0.274382	0.18745	0.15205	0.306799	0.667518
# 16	-0.62756	-0.43653	-0.42416	-0.04608	-0.14311	0.167611	0.187812	0.982884
# 17	0.970119	0.981526	0.951607	-0.23982	-0.48755	-0.23813	-0.38249	-0.47136
# 18	-0.55207	-0.39002	-0.28413	0.584701	0.590399	0.828723	0.828784	0.735967
# 19	0.904194	0.995583	0.993314	-0.19102	-0.375	-0.07373	-0.24393	-0.35794
# 20	0.951711	0.995782	0.979525	-0.23723	-0.40108	-0.18236	-0.339	-0.50623
# 21	0.904194	0.995583	0.993314	-0.19102	-0.375	-0.07373	-0.24393	-0.35794
# 22	-0.45136	-0.31482	-0.23265	0.552625	0.359564	0.739514	0.738787	0.861946
# 23	-0.12985	-0.00111	0.124258	0.527438	0.848124	0.777323	0.722449	-0.07809
# 24	0.946639	0.995187	0.995186	-0.11163	-0.30609	-0.04416	-0.20517	-0.45522
# 25	-0.10177	-0.03797	0.051	0.740706	0.394918	0.841573	0.813879	0.637987
# 26	0.925193	0.998169	0.989615	-0.19906	-0.40399	-0.10866	-0.27381	-0.38088
# 27	-0.67034	-0.61333	-0.53147	0.651755	0.55128	0.744402	0.80665	0.786335
# 28	-0.6441	-0.42823	-0.39622	0.029301	-0.01817	0.287793	0.297072	0.976563
# 29	-0.66295	-0.52839	-0.4306	0.57887	0.598021	0.783291	0.812089	0.763034

	#6	#7	#8	#9	#10	#11	#12	#13
# 30	-0.23774	-0.1905	-0.06792	0.877788	0.789338	0.994654	0.986808	0.392365
# 31	-0.59954	-0.3681	-0.35281	-0.12355	-0.16206	0.145503	0.145766	0.955563
# 32	-0.39437	-0.30965	-0.20029	0.776478	0.646954	0.923485	0.92793	0.641617
# 33	-0.66032	-0.46694	-0.41111	0.238124	0.181541	0.482648	0.497773	0.961825
# 34	0.43811	0.509247	0.556542	-0.06094	0.325089	0.092615	-0.00969	-0.73519
# 35	-0.74279	-0.60824	-0.5868	0.119789	0.023806	0.268215	0.325048	0.978256
# 36	-0.28413	-0.27416	-0.16685	0.567218	0.950835	0.66461	0.681395	-0.23831
# 37	0.661352	0.700044	0.679688	-0.45582	-0.18837	-0.38469	-0.4943	-0.82544
# 38	-0.07281	0.230557	0.331139	0.170746	0.335245	0.605327	0.481447	0.381454
# 39	0.294247	0.368116	0.48407	0.716945	0.612372	0.874015	0.777413	0.016271
# 40	-0.31409	-0.1777	-0.15363	-0.36221	0.30439	-0.14186	-0.15015	-0.35197
# 41	-0.91581	-0.72785	-0.67831	0.06452	0.322543	0.316055	0.373326	0.762853
# 43	-0.25981	-0.29646	-0.17816	0.846908	0.988752	0.877257	0.905318	-0.03126
# 44	0.869275	0.82744	0.858355	0.359206	-0.02303	0.320408	0.19825	-0.30971
# 45	0.763537	0.921444	0.958351	-0.07978	-0.11059	0.148641	-0.03344	-0.32048
# 46	0.837771	0.922404	0.944006	-0.12834	-0.08977	0.010088	-0.1524	-0.59962
# 47	-0.57744	-0.3803	-0.33478	0.183496	0.062197	0.423637	0.426978	0.982333
# 48	0.496484	0.506833	0.606206	0.624321	0.58981	0.705619	0.605935	-0.35334
# 49	-0.53852	-0.50861	-0.40278	0.818045	0.756613	0.898634	0.946162	0.578444
# 50	-0.83182	-0.69053	-0.60194	0.429596	0.627616	0.645853	0.699679	0.686932
# 51	0.736839	0.774533	0.812237	0.008441	0.157864	0.105852	-0.02712	-0.74008
# 52	-0.66031	-0.61348	-0.55287	0.563832	0.378857	0.625539	0.694497	0.868587
# 53	-0.77142	-0.68248	-0.64776	-0.10267	0.489956	0.049945	0.126041	0.0405
# 54	0.367716	0.341349	0.311618	-0.46113	-0.04657	-0.47001	-0.50886	-0.88298
# 55	-0.42921	-0.45186	-0.34582	0.66605	0.985678	0.715048	0.766135	-0.09273
# 56	-0.66904	-0.40326	-0.3446	-0.18412	0.352781	0.193469	0.18271	0.33673

	# 14	# 15	# 16	# 17	# 18	# 19	# 20	# 21
# 14	1							
# 15	-0.9016	1						
# 16	-0.46246	0.633404	1					
# 17	0.98582	-0.85876	-0.47577	1				
# 18	-0.40123	0.454719	0.664848	-0.51598	1			
# 19	0.991438	-0.88989	-0.36047	0.965024	-0.31195	1		
# 20	0.996074	-0.92327	-0.50251	0.98777	-0.46805	0.983118	1	
# 21	0.991438	-0.88989	-0.36047	0.965024	-0.31195	1	0.983118	1
# 22	-0.32197	0.513654	0.775897	-0.40085	0.944445	-0.23421	-0.399	-0.23421
# 23	-0.00785	-0.2158	-0.12333	-0.16858	0.596197	0.032107	-0.03482	0.032107
# 24	0.997432	-0.90516	-0.47271	0.973484	-0.35513	0.989782	0.989913	0.989782
# 25	-0.03012	0.285082	0.501457	-0.10877	0.826924	0.028397	-0.11809	0.028397
# 26	0.996111	-0.88451	-0.38498	0.979128	-0.35722	0.998117	0.989468	0.998117
# 27	-0.61274	0.726861	0.695325	-0.67957	0.928903	-0.55108	-0.67845	-0.55108
# 28	-0.4558	0.593407	0.987686	-0.49497	0.760725	-0.34599	-0.50033	-0.34599
# 29	-0.53787	0.582993	0.695344	-0.63808	0.98747	-0.45557	-0.60013	-0.45557
# 30	-0.18099	0.241091	0.259448	-0.30254	0.872649	-0.13785	-0.25455	-0.13785
# 31	-0.39977	0.537478	0.989939	-0.42538	0.656696	-0.2882	-0.43557	-0.2882
# 32	-0.30709	0.423892	0.524693	-0.41276	0.95821	-0.24238	-0.38478	-0.24238
# 33	-0.48842	0.619796	0.940505	-0.54847	0.878932	-0.38339	-0.54399	-0.38339
# 34	0.502426	-0.80855	-0.69987	0.409675	-0.22929	0.487437	0.533153	0.487437
# 35	-0.62621	0.784109	0.971208	-0.63827	0.726163	-0.53914	-0.66955	-0.53914
# 36	-0.26939	-0.02151	-0.28464	-0.39487	0.439377	-0.27171	-0.27727	-0.27171
# 37	0.692001	-0.91659	-0.74651	0.673614	-0.64581	0.658751	0.745028	0.658751
# 38	0.200266	-0.25124	0.380529	0.0497	0.707435	0.309389	0.157325	0.309389
# 39	0.376657	-0.33145	-0.10385	0.243484	0.569795	0.408248	0.313293	0.408248

	# 14	# 15	# 16	# 17	# 18	# 19	# 20	# 21
# 40	-0.20505	-0.22782	-0.23039	-0.27494	-0.06659	-0.18732	-0.14791	-0.18732
# 41	-0.75589	0.673882	0.799364	-0.81836	0.755978	-0.66469	-0.77493	-0.66469
# 43	-0.28013	0.14753	-0.14203	-0.39844	0.618548	-0.28255	-0.3196	-0.28255
# 44	0.848154	-0.64975	-0.41497	0.816364	-0.09434	0.826767	0.806915	0.826767
# 45	0.910278	-0.9024	-0.32178	0.830337	-0.08395	0.941189	0.896625	0.941189
# 46	0.91833	-0.98495	-0.59146	0.857701	-0.32159	0.913787	0.925443	0.913787
# 47	-0.40209	0.579509	0.962131	-0.45082	0.829639	-0.29547	-0.45945	-0.29547
# 48	0.521757	-0.55705	-0.45962	0.404087	0.26053	0.516032	0.481852	0.516032
# 49	-0.50116	0.561872	0.462341	-0.59553	0.923475	-0.45529	-0.5689	-0.45529
# 50	-0.7059	0.636074	0.660478	-0.80218	0.910937	-0.62746	-0.74583	-0.62746
# 51	0.776412	-0.93693	-0.74237	0.70201	-0.30484	0.752795	0.790262	0.752795
# 52	-0.61288	0.791541	0.783051	-0.65196	0.867354	-0.55124	-0.67856	-0.55124
# 53	-0.70372	0.339109	0.124664	-0.76078	0.262081	-0.6784	-0.66703	-0.6784
# 54	0.336664	-0.66652	-0.78676	0.33712	-0.69242	0.281018	0.409525	0.281018
# 55	-0.44278	0.199224	-0.15738	-0.55485	0.533966	-0.4466	-0.45907	-0.4466
# 56	-0.44337	0.16071	0.43165	-0.54812	0.525319	-0.35278	-0.43166	-0.35278

	# 22	# 23	# 24	# 25	# 26	# 27	# 28	# 29
# 22	1							
# 23	0.34052	1						
# 24	-0.28841	0.056741	1					
# 25	0.913478	0.375779	0.013397	1				
# 26	-0.2682	-0.01612	0.992041	0.003762	1			
# 27	0.927899	0.364748	-0.57533	0.800944	-0.57914	1		
# 28	0.833166	0.032827	-0.45621	0.563214	-0.3785	0.751015	1	
# 29	0.931813	0.532386	-0.49537	0.779653	-0.4968	0.965913	0.779599	1
# 30	0.798758	0.733957	-0.1165	0.871968	-0.17279	0.808357	0.372987	0.836751
# 31	0.746199	-0.07052	-0.41029	0.459532	-0.31808	0.6386	0.987378	0.67357
# 32	0.935865	0.583386	-0.25393	0.925012	-0.27821	0.926171	0.616502	0.939328
# 33	0.923228	0.191254	-0.47403	0.692354	-0.41941	0.86929	0.975909	0.894531
# 34	-0.46726	0.621243	0.530802	-0.32988	0.466296	-0.503	-0.60185	-0.3187
# 35	0.815847	-0.07293	-0.62859	0.544186	-0.56073	0.814305	0.96339	0.781329
# 36	0.148103	0.904215	-0.20832	0.162108	-0.305	0.3282	-0.14973	0.433375
# 37	-0.76829	0.136358	0.684702	-0.62326	0.660011	-0.86755	-0.72225	-0.72894
# 38	0.600109	0.719957	0.239101	0.547173	0.251636	0.397953	0.503278	0.602906
# 39	0.496035	0.775358	0.439503	0.721924	0.375154	0.368942	0.019864	0.459787
# 40	-0.36333	0.473699	-0.19759	-0.56024	-0.21663	-0.24094	-0.15696	-0.05675
# 41	0.665751	0.285467	-0.74444	0.317759	-0.70369	0.75245	0.845918	0.814892
# 43	0.417057	0.849883	-0.21137	0.498606	-0.31078	0.572887	-0.01612	0.611377
# 44	0.005945	0.158672	0.872138	0.38621	0.832493	-0.22357	-0.38646	-0.21983
# 45	-0.10013	0.352317	0.926786	0.123054	0.920418	-0.4	-0.25686	-0.2396
# 46	-0.37679	0.314721	0.931752	-0.11983	0.903758	-0.60669	-0.53574	-0.46164
# 47	0.914245	0.097696	-0.39365	0.700261	-0.32816	0.819852	0.984194	0.836856
# 48	0.146144	0.752013	0.581125	0.440901	0.493651	0.054288	-0.34187	0.147438

	# 22	# 23	# 24	# 25	# 26	# 27	# 28	# 29
# 49	0.863573	0.581221	-0.44716	0.822183	-0.48625	0.952239	0.550193	0.939458
# 50	0.788866	0.541289	-0.66907	0.549883	-0.67012	0.896828	0.744607	0.951396
# 51	-0.44009	0.478851	0.801983	-0.18064	0.741757	-0.5708	-0.6649	-0.42564
# 52	0.927826	0.173346	-0.58745	0.78787	-0.57201	0.97885	0.809587	0.913478
# 53	-0.00497	0.42464	-0.69109	-0.30635	-0.70697	0.235361	0.186632	0.337759
# 54	-0.88023	0.113999	0.326973	-0.82435	0.287914	-0.8216	-0.77117	-0.7145
# 55	0.272045	0.835977	-0.3819	0.258941	-0.47625	0.493868	-0.03474	0.554408
# 56	0.301767	0.545245	-0.42786	-0.01689	-0.40398	0.344893	0.521083	0.533752

	# 30	# 31	# 32	# 33	# 34	# 35	# 36	# 37
# 30	1							
# 31	0.231262	1						
# 32	0.957083	0.490679	1					
# 33	0.561678	0.929454	0.770151	1				
# 34	-0.00129	-0.60314	-0.24003	-0.55454	1			
# 35	0.363738	0.936166	0.613852	0.952392	-0.75235	1		
# 36	0.623428	-0.26648	0.447841	0.021935	0.565683	-0.15499	1	
# 37	-0.4734	-0.6496	-0.66314	-0.77503	0.856215	-0.86777	0.082513	1
# 38	0.597618	0.468085	0.606168	0.55768	0.330567	0.291194	0.376309	0.025682
# 39	0.830243	-0.08494	0.689849	0.186678	0.404456	-0.09056	0.549275	0.033936
# 40	-0.17524	-0.13726	-0.25765	-0.17393	0.647815	-0.24064	0.559617	0.493581
# 41	0.388351	0.806691	0.566948	0.858482	-0.36484	0.848365	0.264025	-0.62822
# 43	0.852349	-0.16568	0.704951	0.193745	0.307402	0.017951	0.914717	-0.20524
# 44	0.259083	-0.40929	0.109111	-0.31708	0.328416	-0.48223	-0.06421	0.342926
# 45	0.07822	-0.22706	-0.0515	-0.25482	0.658132	-0.48704	0.020142	0.66922
# 46	-0.07883	-0.5008	-0.26784	-0.53321	0.798411	-0.73261	0.086053	0.843963
# 47	0.504373	0.951469	0.727625	0.990232	-0.60053	0.950917	-0.10371	-0.76312
# 48	0.631278	-0.43199	0.408187	-0.18059	0.648226	-0.44583	0.598998	0.336085
# 49	0.93522	0.411071	0.966044	0.717696	-0.26223	0.60235	0.555132	-0.71608
# 50	0.700489	0.650517	0.803902	0.840028	-0.24997	0.759722	0.521601	-0.66248
# 51	0.006091	-0.66002	-0.23796	-0.61924	0.922094	-0.8255	0.346699	0.865253
# 52	0.702572	0.71586	0.86412	0.8971	-0.66318	0.887229	0.131925	-0.9397
# 53	0.064586	0.159269	0.072942	0.206806	0.202961	0.210354	0.608401	-0.05679
# 54	-0.54825	-0.71112	-0.73712	-0.81664	0.800036	-0.835	0.232676	0.905067
# 55	0.693064	-0.16715	0.554653	0.147873	0.363299	0.008805	0.973045	-0.13736
# 56	0.214771	0.512981	0.282504	0.516472	0.219776	0.413977	0.477066	-0.0509
	# 30	# 31	# 32	# 33	# 34	# 35	# 36	# 37

	# 38	# 39	# 40	# 41	43	44	45	46
# 38	1							
# 39	0.687714	1						
# 40	0.23198	-0.1402	1					
# 41	0.439689	-0.05852	0.289296	1				
# 43	0.370458	0.708401	0.185274	0.258789	1			
# 44	0.208236	0.681215	-0.49904	-0.70233	0.109225	1		
# 45	0.567234	0.59779	0.031205	-0.47049	-0.03119	0.769661	1	
# 46	0.35237	0.485272	0.151561	-0.64545	-0.03136	0.741752	0.949996	1
# 47	0.541991	0.165021	-0.26506	0.795345	0.086582	-0.25153	-0.20216	-0.49476
# 48	0.501876	0.929614	-0.0044	-0.3382	0.672358	0.753834	0.678584	0.676007
# 49	0.453019	0.579534	-0.1656	0.616836	0.787946	-0.05665	-0.25796	-0.42199
# 50	0.529207	0.271358	0.20398	0.922206	0.594511	-0.47642	-0.38276	-0.55389
# 51	0.295716	0.522209	0.313797	-0.619	0.194056	0.655864	0.840861	0.95164
# 52	0.302929	0.243651	-0.3701	0.734506	0.405861	-0.24482	-0.45961	-0.68262
# 53	0.140059	-0.22567	0.833054	0.671984	0.35903	-0.80327	-0.45064	-0.3884
# 54	-0.20287	-0.20367	0.692812	-0.48004	-0.11684	-0.00396	0.307799	0.558486
# 55	0.298708	0.498854	0.444215	0.372095	0.949753	-0.15705	-0.1715	-0.12539
# 56	0.618438	0.027483	0.750164	0.805815	0.262457	-0.60408	-0.07632	-0.16762

	# 38	# 39	# 40	# 41	43	44	45	46

	47	48	49	50	51	52	53	54
# 47	1							
# 48	-0.20796	1						
# 49	0.648872	0.327718	1					
# 50	0.756371	-0.00165	0.859139	1				
# 51	-0.62062	0.761538	-0.32638	-0.46766	1			
# 52	0.86985	-0.09275	0.87833	0.828917	-0.69373	1		
# 53	0.089088	-0.22953	0.238184	0.606531	-0.19039	0.131477	1	
# 54	-0.84779	0.133965	-0.67969	-0.53933	0.677778	-0.90774	0.268112	1
# 55	0.018221	0.498165	0.688044	0.635667	0.140634	0.316856	0.622189	0.054983
# 56	0.4372	-0.10135	0.309144	0.714161	-0.10036	0.252219	0.827879	0.065637

	47	48	49	50	51	52	53	54

	55	56
# 55	1	
# 56	0.448542	1

	55	56

Vitae

Jenelle Lee has a Bachelor of Arts in Political Science from the University of Iowa, a Master of Science in Public Service Management from DePaul University, an Educational Specialist degree from William Woods University, and expects to receive her Doctor of Education in Educational Administration from Lindenwood University. She holds Missouri certifications in 5 through 12 English and Social Studies and K-8 Building Level Administrator. Scholarly interests include: Grading practices, professional development, school reform, and accountability. Throughout the past 12 years, Jenelle has served as a middle and high school teacher and adjunct instructor. Her prior teaching experiences range grades 5 through 12, dual credit, and college. She currently works as a middle/high school social studies instructor at a rural school in Southeast Missouri, where she resides with her husband Brian, two sons Madden and Archer, and two daughters Whitney and Kenzie.