A QUANTITATIVE STUDY EXPLORING GRADING AND ASSESSMENT PRACTICES IN THE MIDDLE SCHOOL ENVIRONMENT

A Dissertation

Presented in Partial Fulfillment of the Requirements for the

Degree of Doctor of Education

With a

Major in Educational Leadership

in the

Department of Graduate Education

Northwest Nazarene University

by

Greg Wiles

March 2013

Major Professor: Dr. Mike Poe, EdD

AUTHORIZATION TO SUBMIT DISSERTATION

This dissertation of Greg Wiles, submitted for the degree of Doctor of Education with a major in Educational Leadership and titled "Grading and Assessment Practices," has been reviewed in final form. Permission, as indicated by the signatures and dates given below, is now granted to submit final copies.

Major Professor _____ Date 4/15/13
Dr. Mike Poe

Committee Members _____ Date 4-15-13
Dr. Duane Slemmer

_____ Date 4-15-13
Dr. Paul Harman

EdD Program Director _____ Date 4-15-13
Dr. Loredana Werth

Department Administrator _____ Date 5-1-13
Dr. Paula Kellerer

Discipline's College Dean _____ Date 5-1-13
Dr. Paula Kellerer

ACKNOWLEDGMENTS

The journey of completing a doctoral study is not accomplished alone. It involves the contributions of many hearts, hands, and minds, and it is with the utmost gratitude, respect, and appreciation that I acknowledge the following people.

First, thank you to my longsuffering family. I cannot begin to count the number of hours you have sacrificed by allowing me the time and space needed to fulfill the myriad requirements of this program. You tagged along with me on this journey, always providing me with all the love, encouragement, and help I needed to achieve this quest. My office at Lone Star became our second home and you never complained. Thank you, Babe, for reading my dissertation and listening to me babble on about grading and assessment, and qualitative and quantitative research, without falling asleep or throwing a pie in my face. Thank you for bringing the kids to the school so I could at least say hello and goodnight a few times and not feel like the most negligent daddy in the world. Thank you for never making me feel guilty, even though I did. Thank you for making a million delicious home-cooked meals and bringing them to the school for me. Thank you, kiddos, for being content to play basketball, tag, computer games, do homework, read, and whatever else while I was focused on this paper. Thank you all for being so supportive and patient with me. I promise to spend the next several years making it up to you!

Second, thank you to my parents and sisters. Mom and Dad Wiles, thanks for your love, guidance, and direction in the early stages of my educational career. I learned the virtues of hard work and perseverance from you, and I am here because of your sacrificial investment in my life. You are the best parents in the world, and I am honored to be your son. Thank you, Leslie and Denise, for your prayers and happy thoughts and

for checking on me periodically. You are the best sisters and I am so glad God "samiched" me between the two of you!

Third, thank you to my extended family at Lone Star Middle School for supporting and encouraging me through this difficult process. If it were not for you pushing the envelope on grading and assessment, I would not have selected it as my theme. Your personal educational journeys through grading and assessment practices have helped propel this research. I am the luckiest guy alive because I get to go to work every day with teachers who truly love kids and want to see them succeed in learning. You are fun, passionate, and tireless in your efforts to impact our students! I am sorry for the tough financial year, and I hope we have brighter days ahead. Thank you, Aaron Moiso, for asking those difficult questions about grading and assessment years ago. I appreciate you for letting your creative juices flow and spill over onto me. Thanks to you and the rest of the grading and assessment team—Becky Johnson, Shayla McIntyre, Monique Gafford, and Kevin Murphy—for investing your skills and for being so dedicated in the training and preparation it took to plow the way for best practices in grading and assessment. You rock! Thank you, Kim Foster, for generously sharing your superb editing skills, sharp eye, and knowledge of APA with me—no amount of movie passes could ever reimburse you for your time. You and Chuck have been a tremendous blessing to us and our children (as surrogate grandparents), and I will thank you forever for saving me from the bibliography meltdown in my office. Thank you, Linda Goodwin, for the many hours of counsel, wisdom, and support over the years. You are a true master teacher and an amazing instructional coach. I sure miss having you at my fingertips at Lone Star, but I know you are enjoying retirement—you deserve it. I promise not to

bother you ever again with a research project! Thank you, Greg Heideman, for being my right-hand man, my technological support, and my loyal friend. You have rescued me from computer catastrophes and helped me maintain some semblance of sanity. Thanks to you and Joy for breathing life and friendship into Des and me. You are more than friends—you are priceless treasures to us. Thank you, Randy Jensen, for your never-ending support and encouragement throughout my educational journey. You were the best mentor I could have asked for, and I learned so much from you about servant leadership and having fun along the way. Jeff Read, you are also my best mentor, and your friendship has been one of the most invaluable rewards of choosing this career. Thank you for believing in me and taking a chance on a young, energetic, P. E. and Health guy many years ago. Thanks also to Chad Rawlins, my "stats man," for talking to me in my own vernacular instead of statistical jargon—you made it all much less complicated when you put it into colors and critters.

Thank you to the staff and administration of Fall River Middle School (pseudonym) for allowing me to conduct the research on your campus. You are an educational testament of what a true middle school should look like. Thank you for looking inward throughout this process and for making strides as a staff toward better grading and assessment practices.

I have been blessed to experience this process with a great cohort of doctoral students. I could not have continued to endure the suffering without knowing you were suffering, too. Misery loves company, you know—thanks for being my company! Thank you, especially, Jim Brown, Heidi Curtis, and Kelsey Williams, my close comrades these past few years—we made it! Thank you all for being available, even late at night, and for

talking me off the ledge numerous times. Thank you for helping me find the humor in the agony. If we ever do something like this again, I would want you on my side for support (or to shoot me in the head). Jim, sorry our high hopes for working out during classes came crashing down when the assignments kept piling up. Midnight dinners and sleep deprivation took a toll on us. Enjoy getting back to motorcycle riding, exercise, and rekindling the romance with Kathy. Heidi, you will make a fantastic addition to a doctoral staff someday. Enjoy Hawaii and Washington DC, and don't be surprised if we follow you to Maui! Kelsey, enjoy breathing and being with your family again—and not dreaming about dissertation research!

Finally, I want to thank my dissertation committee: Dr. Poe, Dr. Slemmer, Dr. Harman, and the NNU staff (Dr. Kellerer, Dr. Werth and Cindy Williams) for not giving up on me and for encouraging me when I became discouraged. Thank you, Dr. Poe, for talking me into the program and for being there for me, even during odd hours. You continually provided me with timely inspiration to keep going. Thank you, Dr. Slemmer, for challenging the structure of my writing—I know it was for my good, and I hope it reads less like the book of Proverbs and more like a research paper. Thank you, Dr. Harman, for the memorable advice about eating the cow. I have embraced it by taking many little bites, and I have quoted you more than once. Thank you, Dr. Kellerer, for your immeasurable support and encouragement along the way and for helping me initially break down my topic into understandable headings. Nampa School District's loss was NNU's gain—you left a huge hole that is yet to be filled, but it is great to see the impact you are having on higher education. I never stopped liking you, but I will smile bigger every time I see you now. Thank you, Dr. Werth, aka "The Machine," for

expecting more from me than I thought I could give. No one can keep up with you but I am sure it was painfully amusing to watch us all try.

DEDICATION

This study is dedicated to my beautiful and enduring wife Desiree, my pensive son Justus, my compassionate daughter Avery, my cheerful daughter Liberty, and my hilarious son Donovan. You are all my dreams come true. Thank you for your unconditional love and support. I hope this dissertation is an example of what I believe—with God anything is possible (Matthew 19:26)! I am proud of you and eternally grateful to be part of TEAM WILES. I love you with all my heart and pray God is always at the forefront of your lives. He has great plans and an amazing future for each of you (Jeremiah 29:11), so be strong and courageous and He will be with you wherever you go (Joshua 1:9). Give everything you have and are for Him and help others along the way. God has blessed us to be a blessing!

ABSTRACT

Educators are committed to providing students the best possible education by using the practices and tools they know (Zmuda, Kuklis, & Kline, 2004). Unfortunately, the status quo for grading and assessment practices is faulty, outdated, and ineffective. Current practices have unwittingly hindered student progress and motivation (O'Connor, 2007). Educational leaders, administrators, and teachers are faced with questions regarding the best ways to motivate students and accurately report their progress (Popham, 2011). Assessing and grading student achievement are primary functions of educators and if these practices are not approached and applied properly, a grade can misrepresent the true knowledge of the student regarding class standards (Brookhart, 2009). Even though there has been a shift to standards based teaching, grading and assessment practices have not followed suit and continue to lag behind (Zmuda, Kuklis, & Kline, 2004). Conventional grading and assessment practices abound, yet opportunities for the introduction of reform and the necessary professional development for the implementation and sustainability of such reform are lacking (Wormeli, 2006). Second order change is an apparent break from the past and is best presented and maintained through quality, ongoing professional development (Marzano, 2003). Change is inevitable and necessary at times, offering fresh perspective and new ideas, and it is the responsibility of those in the teaching profession to continue to learn and stay apprised of current research to give students the best possible education (Schimmer, 2012). When a district or school becomes aware of the need for change, there is an obligation to address the underlying issues and all relevant facts, logic, and research. Professional development provides the optimal venue for learning, discussing, applying, and evaluating any proposed changes (DuFour &

Marzano, 2011). Reeves (2007) asked, “What is the risk if we engage in this change compared to the risk of continuing our present practice” (p. 7). The risk of embracing a proposed change may include limited buy-in, fear of failure, or the perception of more work. Regardless of potential effects, growth can only come through change, and trying new methods is the only way to know what changes work (Schimmer, 2012). The continuation of current grading practices poses a threat to the educational process and the future of student learning by undermining motivation and presenting inaccurate information due to grades’ convoluted nature (Brookhart, 2009). Grades should not represent tradition, opinion, behavior, effort, homework, attendance, or any other extra-curricular components (O’Connor, 2007). They should purely and simply report the student’s comprehension of predetermined class standards and goals (Brookhart, 2009). There are new research based methods for grading and assessment that honor the meaning of the grade and could be introduced and implemented in schools across the country through professional development opportunities (Reeves, 2007). The purpose of this study was to investigate the perceptions of grading and assessment practices in Fall River Middle School and to analyze the impact of increased opportunities for professional development in grading and assessment practices on the staff.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

LIST OF FIGURES

Chapter I

Grading is a vehicle in education that provides a specific form of communication regarding student progress, but it is questionable whether it accurately meets the needs of all stakeholders (Davies, 2009). Grades are difficult to interpret because teachers use subjectivity and consider many factors, such as behavior and effort, when assigning grades (Brookhart, 2009; Guskey, 2004). Formal training in grading methods has been limited due to the lack of available of resources and applicable professional development (Brookhart, 2009; Guskey, 2004; Shanahan, 2011). The objective of sound educational assessment is to create a method for grading in which grade reports would only reflect achievement on standards, while reserving a separate process for assessing and reporting student behavior, attendance, effort, attitude, and other nonacademic issues (Brookhart, 2011; O'Connor, 2007; Yesbeck; 2011). Professional development is a key element to achieving this goal (Schimmer, 2012).

Statement of the Problem

Professional development and training in best practices for grading and assessment are uncommon. Many teachers find grading a difficult subject, yet it is a primary educational obligation (Guskey, 2004; Yesbeck, 2011). Grading methods have long been a teacher's personal practice or tradition, rather than a school or district policy based on researched practices (Erickson, 2006; Guskey, 2011; Jung & Guskey, 2011; Martinez, Stecher, & Borko, 2009; O'Connor & Wormeli, 2011). In many schools and districts across the country, grading procedures are left in the hands of individual teachers and departments (Erickson, 2011; Guskey, 2011; Guskey & Jung, 2012). Practices such as recording zeros are widely employed but do not sufficiently reflect a student's

progress and, furthermore, may present inaccurate evidence and could falsely represent learning deficiencies (Brookhart, 2011; Guskey & Bailey, 2001). Unfortunately, zeros are commonplace in most grade books. Many district policies lack an overall consistency and congruency between and among departments and grade levels (Allen, 2005; Brookhart, 2009; Brookhart, 2011; Guskey & Bailey, 2001), which leads to imprecise findings, poor communication, and futile measures (Daniels, 2010; O'Connor, 2007; Schimmer, 2012; Sousa, 2006; Sylwester, 2007; Zmuda, 2010). Educators have not come to an agreement on grading practices, and the controversy over change has induced heated debates regarding the separation between academic and nonacademic factors (Brookhart, 2011; Reeves, 2011). Schools are complex systems, and when one part is dysfunctional, the entire structure becomes out of alignment (Allen, 2005; Brookhart, 2011; Reeves, 2011).

Background

Despite recent pressure in education to achieve certain standards, conventional practices in grading and assessment continue, particularly in middle schools and high schools (O'Connor, 2007; Schimmer, 2012). New research and requirements make it necessary to reexamine current methods for their worthiness in consistency and accuracy (Hess, 2010). Adequate, proper training has not been made available to most teachers, and consequently, they are not prepared with the best methods and tools in grading and assessment practices (Chappuis, Stiggins, Chappuis, & Arter, 2012; Guskey, 2004; Moss & Brookhart, 2009; Yesbeck, 2011). What exists in current assessment training is too varied and has been concentrated on the state and federal mandates for assessment rather than the regular reporting of progress on standards in the classroom. Keeping up with educational changes and directives, including the No Child Left Behind Act of 2001, has

been the primary focus of educators the past few years (Schimmer, 2012; Yesbeck, 2011).

Research Question

The central research question for this study was: Can grading and assessment perceptions among middle school educators change through professional development?

Description of Terms

- **Assessment:** the process of collecting enough data pertaining to a student learning to ascertain how the student is achieving determined standards. Assessment has two parts: assessment for learning (formative) and assessment of learning (summative); may also refer to organized activities to establish a student's knowledge or skill base in a given capacity.
- **Formative assessment:** the process of gathering information throughout a student's learning and involves useful feedback, including homework, class activities, group practice, rough drafts, lab activities, and quizzes.
- **Summative assessment:** the tool for measuring a student's achievement on standards, allowing them to demonstrate their knowledge after sufficient time has been allotted for preparation. Summative assessments may include tests, projects, writing tasks, reflections, and lab assessments.
- **Evaluation:** the process of measuring the quality of work based on identified criteria and assigning value to signify the level of achievement reached.
- **Measurement:** the allocation of grades or scores to an assessment based on a certain set of standards.

- **Test:** an assessment that takes place at a given time and most often uses a pencil and paper format.
- **Feedback:** constructive comments and advice pertaining to a student's work; it is frequent and specific in order to be beneficial and increase student understanding.
- **Standardized test:** an assessment by which students take the same test under the same conditions with the same directions and grading system.
- **Large-scale assessment:** a test given to a large number of students within classrooms and schools at the same time.
- **Grade:** the symbol assigned at the end of a session (e.g., quarter or semester) as a summary report of a student's performance or achievement level.
- **Score:** the letter or number given based on a process of assessment.
- **Mark:** another name for a score or grade and may also refer to a comment or recommendation in the spirit of formative assessment.
- **Standards-based grading:** state targets by which students' achievements are measured.
- **Purposeful community:** a group of people dedicated to working together, believing they can make a difference by achieving a common goal.
- **Common Core State Standards:** college and career ready common benchmarks shared by many states in Math, English Language Arts, Literacy in History/Social Studies, Science and Technical Subjects.
- **Professional learning communities:** collaborative groups researching and applying common practices to improve student achievement.
- **Professional development:** strategic time set aside for learning.

Significance of the Study

Perceptions in grading and assessment may change with strategic, research-based professional development opportunities, such as best practices grading conferences and book studies (Dyb, 2012). Too often, professional development becomes a "one-hit wonder" and is forgotten or not implemented into the day-to-day operations of a school (Wormeli, 2012). Professional development should be entrenched in a school's philosophy and practice. It sets the stage for building a purposeful community, a middle level administrator's primary objective (Eaker & Keating, 2012). Within the purposeful community, changes and new ideas are presented, discussed, and implemented to serve and educate students with the most current, successful, and research-based methods available (Dufour & Marzano, 2011). An environment is established where the staff performs daily routines, such as grading and assessment practices that are fair and consistent and have been previously agreed upon through a collaborative process and professional development (Erkens & Twadell, 2012). In order to sustain change, a staff must participate and benefit from ongoing professional development (Allen, 2005; Brookhart, 2011; Dyb, 2012; Erickson, 2011; Fisher, Frey, & Pumpian, 2011; Schimmer, 2012).

This study aimed to challenge the status quo regarding grading and assessment practices by presenting professional development opportunities to the staff of Fall River Middle School. Prior to the professional development sessions, a presurvey was administered to ascertain perceptions of the participants. Next, professional development was presented through one in-service training consisting of five classes, and six follow-up lessons to be taught by FRMS administration and staff. A postsurvey was conducted

following the professional development and the data was analyzed to determine whether the sessions made an impact on teachers' perceptions. The results of this study indicated a positive shift in perceptions in most of the survey questions, after the limited professional development. This study helped to reinforce the research pointing to the need and benefits of professional development.

Overview of Research Methods

This study of best practices in grading and assessment among middle school educators was implemented at Fall River Middle School (FRMS) located in a small rural community in the northwestern part of the United States. FRMS identified a grading and assessment theme for the school year (2012–2013) as its professional development emphasis. All teachers at FRMS participated in the in-service training for grading and assessment within their professional learning communities (PLCs) as part of the regularly scheduled professional development modules. This in-service training was scheduled through the principal at the beginning of the school's year-long professional development calendar. The teachers were required to attend professional development sessions throughout the year and were formally introduced to the theme and outline, then invited to participate in the survey portion of the study during the in-service training. The collaborative climate of FRMS was conducive to participation in this study through the PLC setting. The Informed Consent form issued through Qualtrics, clearly stated participants were able to withdraw from the study at any time.

The researcher guest conducted the research as a quantitative study using a presurvey and postsurvey (see Appendices A & B) during the fall semester of 2012. All

identifying information was deleted from this report in accordance with the Family Education Rights and Privacy Act (20 U. S. C. § 1232g: 34 CFR Part 99).

Chapter II

Review of Literature

The Need for Change

For years the educational system has used letter symbols and numeric point values to calculate a grade meant to reflect a student's level of knowledge and understanding of a particular subject. This method of reporting has been challenged because it provides an insignificant contribution to the educational process (Tomlinson, 2011), and a letter or numeral may represent many different components besides the student's actual comprehension of subject matter. The grade is often confusing to parents, inconclusive to administrators, and potentially discouraging to students (Brookhart, 2011; Schimmer, 2012). As long as schools continue this practice, it poses a dilemma because the grade is not pure, but rather reflects numerous variables impossible to differentiate and clearly understand (Carey, 2001; Stiggins & Knight, 1998).

The complexity of a grade on a typical report card makes understanding difficult and accurate assessment unattainable. A grade may incorporate aptitude, achievement, effort, compliance, and attitude within a single number or letter (McMillan, Myran, & Workman, 2002; McDaniel, 2010; Shanahan, 2011; Stiggins, 2001). Aptitude could indicate that those who are naturally able to learn more throughout the course receive a better grade (Stiggins, 2001). Achievement grades may be given to students who go above and beyond the expectation of the class (Stiggins, 2001). Students who are measured according to effort by working harder than their peers could end up with a higher mark than their true knowledge warrants. Compliant students who follow classroom rules may also be given higher grades. Finally, a high grade for attitude could

be given to students who are positive and have good manners (Kohn, 2011; McMillan, Myran, & Workman, 2002; Stiggins, 2001). A grade loaded with extracurricular elements does not truthfully convey an academic measurement and hinders the educational process (Brookhart, 2011; O'Connor, 2007).

It is too easy for teachers to pollute grades when they include assessments of student effort, attitude, and other contributors within the grade (Carey, 2001; Stiggins & Knight, 1998). Teachers (including those who regularly use formative and summative assessments) continue to include nonacademic elements to calculate grades (Goodwin, 2011; Stephens, 2010). Even though there is a standards-based curriculum, grading practices continue to represent outdated philosophies (Guskey & Jung, 2012), and teachers are still grading attitude, effort, and other factors in combination with measurements of student achievement on academic standards (Duncan & Noonan, 2007; Stiggins & Knight, 1998). One grade represents several elements, such as attitude, effort, behavior, homework, and participation, in addition to the achievement level (Aronson, 2008; Brookhart, 1994; Brookhart, 2011; Hummel, 2011; McMillan & Nash, 2000; McMillan, Myran, & Workman, 2002; O'Connor, 2007; Stephens, 2010). Current and past grading practices have also included marks for attendance and cheating within a single letter or number value (Annerstedt & Larsson, 2010; Fleenor, Lamb, Anton, Stinson, & Donen, 2011; Scherer, 2011). When teachers report students' grades containing more facets than the real knowledge of the subject matter, it clouds the authenticity of academic assessment (Carey, 2001; Stiggins & Knight, 1998).

It is crucial to differentiate the components of a grade, instead of lumping them all together, so academic information is uncomplicated and clearly reported (Goodwin,

2011; O'Connor, 2007). Stiggins and Knight (1998) pointed out the importance and necessity "to indicate each student's current level of academic achievement with nothing else factored in to interfere with that message" (p. 61). When academic achievement is the only factor indicated in the letter grade, it provides a clearer snapshot of the child's progress and level of knowledge (Carey, 2001; Stiggins & Knight, 1998). Academic achievement and progress must be reported apart from nonacademic issues to guarantee an accurate assessment of scholastic understanding (O'Connor, 2007; Roorda, 2008), and anything not related to the achievement level (such as work turned in late or extra credit) should not be recorded in the achievement section of the report card (Arter & Chappuis, 2006; Dyb, 2012; McDaniel, 2010; O'Connor, 2007; Roorda, 2008). Other components may still be addressed, although separate from the academic report card (Allen, 2005; Stiggins & Knight, 1998).

Schools must decide upon a method for recording and representing grades in a report card by first defining the purpose of a grade. They need to keep in mind the importance of separating students' achievement of standards from everything else. The essence of the grade must reflect a student's knowledge of concepts and the goal for maintaining a grade's purity should not be negotiable (Guskey, 2011; Scherer, 2011; Stiggins, 2001). It is vital for academic grades to be separated from effort and behavior, and teachers and administration can still require a high level of responsibility on the part of students (Arter & Chappuis, 2006; O'Connor, 2007; Haselhuhn, Al-Mabuk, Gabriele, Groen, & Galloway, 2007).While behavior issues are important and need to be communicated with parents and students, recorded grades should only reflect a student's knowledge of concepts and academic progress (Balfanz, 2009; Grimes, 2010; Marzano &

Heflebower, 2011; Rom, 2011). The needs and benefits of recording behavior and engagement should be assessed apart from the academic grades to produce an authentic report card (Brookhart, 2011; Davies, 2009). For a grade to be a reliable source of information, it must be a true and just representation of a student's knowledge, separate from all other supplementary factors (Scherer, 2011).

Grades exemplify teachers' central philosophies about education and its purposes (Gullen, Gullen, & Erickson-Guy, 2012), and there is a wide spectrum of beliefs on grading practices (Scherer, 2011). The interaction of teachers with colleagues plays a part in influencing their assessment practices, along with pressures for improving student outcomes (Gordon & Fay, 2010; Maloley, 2008). Some educators believe current practices accurately reflect how the real world works and operates, and they are hesitant to discontinue a holistic grading approach because of the ardent belief that students must be held accountable for their choices (Balfanz, 2009, Erickson, 2001). They assert assessing effort, timeliness, and other factors promote education and work ethic (Scherer, 2011). Educators may also express resistance to change because a large number of current practices have deep roots in educational traditions (Guskey & Bailey, 2001; Guskey, 2006). Understandably, teachers' grading philosophies originate from their own history and personal educational experiences, thus teachers are more comfortable with past methods than contemporary or progressive approaches (Guskey & Bailey, 2001; Guskey, 2006). The opinions and judgments of teachers play a big role when delivering course grades, and the conventional practice of recording typical letter grades for student work reflects a reliance on professional judgment based on little objectivity and too much subjectivity (Aronson, 2008; Guskey & Bailey, 2001; Guskey, 2006).

Proposals for change face the existence of two mindsets; people demonstrate what they believe their capabilities are according to a fixed mindset or a growth mindset (Dweck, 2006; DuFour, DuFour, Eaker, & Karhanek, 2010). A fixed mindset is the belief a person has a finite capacity or ability (Dweck, 2006; DuFour, DuFour, Eaker, & Karhanek, 2010; Kohn, 2011; Schenck, 2003). For example, those with a fixed mindset may have been told they were smart by a teacher at an early age; therefore, they always believe they are smart. When a student meets challenges, in order to preserve his fixed self-image, he becomes frustrated and has the propensity to quit (Chappuis, Commodore, & Stiggins, 2010; DuFour, DuFour, Eaker, & Karhanek, 2010; Dweck, 2006). The person with a growth mindset has the belief he can grow and excel with effort and eventually achieve the level of success he desires (Dweck, 2006).

Positive and negative words, along with grading procedures, have an impact on the mindset and the subsequent effort or desire on the part of a student (Dweck, 2006; Kohn, 2011). Marzano, Pickering, and Helfebower (2011) stated “self-efficacy is quite possibly the most important factor affecting engagement” (p. 117). Self-efficacy contributes to a student’s sense of self-worth because it is the internal drive to be effective by producing a desired result. Continual failure through poor grades adversely affects a child’s self-worth, thus the overall academic experience, and becomes a point of reference for life (Chappuis, Commodore, & Stiggins, 2010; Dweck, 2006; Kohn, 2011). Grades have the power to influence how students feel about themselves and their future ability to learn (Dyrness & Dyrness, 2008; Schimmer, 2012). When students are given a failing label, they become susceptible to an attitude of defeat and consequently pull back

from the educational arena. Rather than increasing desire and effort, the failure impedes them (McMillen, 2012).

Inherent Problems in Current Grading Practice

Educators believe motivation asserts significant influence on student learning (Haselhuhn, Al-Mabuk, Gabriele, Groen, & Galloway, 2007; Rieg, 2007). Chappuis and Chappuis (2006) described grades as one of many external ways schools use to motivate students. Grades have been used for years as an external (extrinsic) strategy to motivate or punish (Chappuis & Chappuis, 2006; Goodwin, 2011; Sadler, 2010). Even teachers with the best of intentions have much learning to do with regard to research on student motivation (Schimmer, 2012; Wormeli, 2012). Some students do not take the learning process seriously or show self-motivation unless they know their participation will directly impact the summative grade (Sadler, 2010).

Motivation is a concern for middle school teachers in particular because student motivation tends to decline during the transition from elementary to middle school (Haselhuhn, Al-Mabuk, Gabriele, Groen, & Galloway, 2007; Mainkar, 2008). The middle years are notorious for distinctive and sometimes awkward changes in adolescents as they navigate their physical, emotional, social, and cognitive development (Daniels, 2010; Pridham & Deed, 2012; Sousa, 2006; Wilcox & Angelis, 2009). Adjusting to different teachers, styles, routines, and environments can be challenging when simultaneously attempting to handle drastic personal transformations, so anything increasing student motivation during this unique time is beneficial (Daniels, 2010; Wilcox & Angelis, 2009).

Grading practices have the capability of affecting student motivation, effort, and achievement (Bonesronning, 2004; Guskey, 2011). Students who experience success through good grades demonstrate the desire to continue with the related behavior or achievement with the promises of reward, while students who receive low grade after low grade tend to lose the motivation to try harder on subsequent assignments and tests (Chappuis & Chappuis, 2006; Guskey, 2011). Students buried in a deficit due to late assignments, behavior, or missing assignments may show little to no motivation or enthusiasm specifically in their scores and ultimately in their education (Arter & Chappuis, 2006; Chappuis, 2009; Daniels, 2010).

A zero has a negative effect on a student's motivation. For some students, the threat of receiving a zero does not motivate them to change their habits. Some would rather take a zero so they do not have to complete the assignment (O'Connor, 2007; O'Connor & Wormeli, 2011). A vast deficit reflected in a grade is daunting and makes it seem impossible for the student to be able to raise the grade and pass the course, therefore decreasing a student's motivation (Guskey, 2004). Students recoil and become frustrated when they feel someone is always waiting to catch them being wrong rather than celebrating when they are right (Sousa, 2006). Essentially, a setting based on punitive response can negatively impact students' perceptions of education, and consequently, diminish their motivation for learning (James-Ward, Frey, & Fisher, 2012).

A zero score is an invalid reporting method and does not accurately describe a student's knowledge or comprehension of subject material (Marzano, 2006; O'Connor & Wormeli, 2011). Many school districts employ policies requiring scores of zero, not only as a record of a specific failing grade, but also for late or missing work or when a student

cheats on an assignment (Wormeli, 2006). Zeros are often recorded in a grade book to convey to parents their child shows a lack of responsibility by not completing the assigned work, and they are used as a punitive consequence for a perceived lack of care or a dismissive attitude rather than an indicator of a student's understanding of the subject matter or academic achievement (Guskey, 2004; Schimmer, 2012). Punishing students, who are otherwise proficient, with zeros for failure to turn in homework or turning in late homework represents an ineffective grading practice (Erickson, 2001; Wormeli, 2006) and distorts the point of reference, which should be a student's level of understanding of a particular objective or standard (O'Connor, 2007). Some students may turn in half of the assigned homework yet fully understand the concept and demonstrate proficiency on the summative assessment; however, if they are continually graded on homework, they are technically failing the class (Wormeli, 2006). The student who fails to turn in homework for various reasons will predictably and quickly drop behind the rest of the class when zeros are used as grades (Wormeli, 2006).

Another drawback to recording zeros in a grade book is it signifies an improper numerical value, making it very difficult for a student to pass the course (Brookhart, 2011). A grade is skewed and the true measurement of a student's knowledge is unattainable when zeros are used (Wormeli, 2006). An A, which is 90%–100%, has a much smaller range in scale compared to an F, which is 0%–59%. Assigning an F grade reaps the same ill effects and should be considered an inappropriate measure for grading (Gerke, 2007; Schimmer, 2012). The zero does not effectively assess or transmit information about the student's actual knowledge of what is being taught in class, and it also has the propensity to discourage students from attempting to recover and improve

their grades (Guskey, 2004; Wormeli, 2006). It does not reflect a student's progress, and furthermore, may present inaccurate evidence, possibly falsely representing learning deficiencies (O'Connor, 2007; O'Connor & Wormeli, 2011). Marzano (2006) was firm in his beliefs and research regarding systems of grading currently used and said a teacher should never record a zero as a grade for not turning in work, turning in assignments late, or missing a test. According to Wormeli (2006), it is recommended to "avoid recording zeros for work not done" (p. 129).

Inequitable Grading

When individual schools and districts do not agree on a uniform grading philosophy, they perpetuate inconsistency throughout the program (Guskey & Jung, 2012; Marzano, 2010). Grading can be very subjective, neither uniform nor reliable, and can communicate confusing messages to parents and students (Brookhart, 2011; Davies, 2009; Guskey, 2011; Stiggins, 2001). There is a notable variance in the perception and interpretation of teachers regarding the meaning and purpose of grades; they consider achievement and nonachievement factors differently (Brookhart, 1994; Guskey, 2011; Maloley, 2008). Grades mean different things to different teachers and are consequently not a reliable source of information to students, parents, other teachers, or administrators (Roorda, 2008; Stiggins, 2001). Teachers often define each of the contributing factors within a grade individually and weigh them differently than their counterparts across the hall (Dyb, 2012; McMillan, Myran, & Workman, 2002; Roorda, 2008; Stiggins, 2001).

Varying grading methods are commonly found among schools, subjects, and teachers (Hummel, 2011; Stiggins, 2001). Examples include Science teachers who allocate grades differently than Math and Language teachers, giving higher weight to

effort due to the amount of time and application needed for developing hypotheses and conducting experiments (Resh, 2009). Students who move from class to class in the middle and high school years find teachers' individual grading practices confusing (Gerke, 2007; Stiggins, 2001). It is no wonder scores differ from classroom to classroom because some teachers measure student achievement by standardized test scores more than others (Martinez, Stecher, & Borko, 2009; Proulx, Spencer-May, & Westerberg, 2012). Even though standards are frequently used for some course objectives, there is little reference to standards when it comes to the actual reporting of student grades (Stephens, 2010).

There are substantial inconsistencies in grading from school to school and between city districts; municipal schools show lower scores than independent schools (Annerstedt & Larsson, 2010). Grading in low-income schools is too often used to punish and reward and may even unwittingly convey the ultimate goal is good behavior rather than academic achievement (Goodwin, 2011). There are data sources showing more discrepancies in grading practices amongst teachers in middle school due to the inclusion of behavior than with elementary or high school teachers (McDaniel, 2010). Secondary teachers tend to base grading practices on what would prepare students for college and are more committed to the mathematical calculation of the grade than primary teachers (Guskey, 2009), while elementary teachers use more informal evidence and observation (Brookhart, 1994). Some districts implement grading policies, while others choose to leave it entirely to the discretion of the individual teacher, and even within districts that have applied a policy, there are teachers who find it unclear or difficult to understand (Brookhart, 2011; Guskey & Bailey, 2001; Guskey, Swan, & Jung, 2011). Erratic and

unsubstantiated practices for grading and assessing student knowledge contribute to the confusion regarding the composition of a grade. There will continue to be confusion and misunderstanding of the breakdown of a grade if there is not a homogenous approach to the interpretation and delivery of assessments and report cards (Marzano, 2010).

Regrettably, flawed, inadequate, and outdated grading practices are still a part of the educational system. Some teachers use grading practices allowing students to fail a course, hoping it will teach the students a lesson (Dyrness & Dyrness, 2008; Erickson, 2011). Another practice is to establish an A grade expectation for an assignment, and then mark every submission down from there, making comments about what is wrong with the student's work rather than pointing out any positives (Dyrness & Dyrness, 2008). Many tests have a defective design and do not accurately assess a student's knowledge of subject matter and may adversely impact the final class grade (Salend, 2011). Some teachers admit to determining a set number of A's to be given prior to the beginning of a session. They base grades on the determination and do not exceed it despite student submissions (Scherer, 2011). There is also evidence of extreme cases of poor grading practices, such as purchasing grade points through participation in fundraising campaigns and increasing grades for extramural activities. Examples include giving extra credit points for bringing in the most cans in a food drive, not using all of the allotted bathroom passes in a session (Erickson, 2011), and wearing costumes to school on a special day (Scherer, 2011). Teachers who make homework represent a high percentage of the final grade set up many students for discouragement and frustration, and the final grade does not ultimately report the student's actual knowledge of the material (Smith, 2012). Such measures are still a part of the status quo, and the determination and account of a

student's true knowledge and understanding of subject material continue to be elusive (O'Connor & Wormeli, 2011).

Best Practices and Strategies

While teacher practices and classroom policies are important to student sense-making and engagement in school, it is the school-level structures, policies, and practices that have the potential for the greatest impact on student sense-making and responses to early failure (Farrington, 2008). Teachers have become increasingly disappointed with the current grading practices, recording systems, and processes because of the amount of time involved to monitor and report (Guskey & Bailey, 2001; Guskey, Swan, & Jung, 2011). The absence of mandates is not the issue. In fact, the proliferation of mandates both at the district and building level often overlap and compete with one another, leaving frustrated teachers to use their own discretion to determine how to respond to proposed change (Maloley, 2008).

Educators are desirous to improve student learning and implement measures to provide true assessment (Carey & Carifio, 2011; Davies, 2009). Therefore, they must be willing to honestly examine and scrutinize the purpose and productivity of a practice in order to embrace changes promoting the success of students (Marzano, 2001; Schimmer, 2012). Schools open to experimenting with new approaches in grading demonstrate a commitment to reducing dropout rates and retention for their students and supporting their success (Carey & Carifio, 2011). Educators are more accountable than ever for assessing and reporting student knowledge based on the mastery of predetermined standards (Garrison, Chandler, & Ehringhaus, 2009). Danielson (2002) and Carey (2001)

said a school's grading policy has the greatest potential to affect the future of a student, positively or negatively, transferring far beyond the walls of the school.

Johnson (2013) lists reporting grades in a consistent and timely manner as one of the top five professional responsibilities of teachers. Grades are intended to assess and promote learning, and there are many ways to calculate and report them fairly and accurately (McMunn, Schenck, & McColskey, 2003; Popham, 2008; Winger, 2005). Grading practices serve other functions as well: communicating with parents, informing potential employers, career planning, guiding administrative decisions, and giving promotions and honors (Carlson, 2003; McMunn, Schenck, & McColskey, 2003; McMillan, Myran, & Workman, 2002). Class standards, the teacher's judgment, and assessments should be used to calculate a final grade (Van de Walle, 2001).

Grading and assessment cannot be separated because the focus is to grasp the full scope of the influence they have on a student's education and success (Dahlgren, Fejes, Abrandt-Dahlgren, & Trowald, 2009). Surrendering the practice of assigning zeros as a form of punishment is the first step toward positive change in grading and assessment (Brookhart, 2009; Wormeli, 2006). Standards-based grading is the most recent method to ensure accurate assessment and reports of student knowledge and achievement (Proulx, Spencer-May, & Westerberg, 2012). In addition, a rubric scale ranging from one to four can replace the traditional 100-point scale. On a 100-point scale, work not turned in often reflects a bottom line score of a zero; on a 1-point to 4-point scale, an equal distribution score is given (Wormeli, 2006). An F grade on a current standard report card is anywhere between 0% and 59%, where on the rubric scale, the spread is only four points, and a teacher is not able to give a negative score for missing assignments (Wormeli, 2006).

Also, those who embrace the suggestion to establish a bottom line of 60, as opposed to the bottom line of zero, believe it is more effective and mathematically sound (Schimmer, 2012; Wormeli, 2006). These are among the various grading options proven to be more favorable for students, parents, and teachers than current and past practices (Erickson, 2001; O'Connor, 2007).

In addition to changing the methods of calculating and reporting grades, administering regular assessments is the primary recommended tool for reform in student evaluation (Stiggins, 2007). Assessment is an essential piece to the educational puzzle. While assessment was used in the past to assign students to a college or vocational track, it is now considered a crucial means for imparting and expounding on information to increase student knowledge and potential (Garrison, Chandler, & Ehringhaus, 2009). Future goals can be achieved when students are regularly and accurately assessed on ability and comprehension of predetermined standards (Greenstein, 2012). Assessment should describe a student's understanding and focus less on the assignment of a letter or number (Robinson, 2012). Presenting information, assessing knowledge, and adjusting strategies are threads composing the educational process, and grading is just a small part of the process (Tomlinson, 2011). Assessment centers on the overall knowledge, grasp, and proficiency of the material and process as opposed to the recitation of facts (Wiggins & McTighe, 2011).

Literature strongly supports the use of formative assessment (assessment for learning), a practice that occurs during the learning period, providing quality and immediate feedback to both the student and the teacher (Brookhart, 2009; Brookhart, 2010; Chappuis, Stiggins, Chappuis, & Arter, 2012; McMunn, Schenck, & McColskey,

2003). Formative assessment is an invaluable tool and offers the most benefits for teachers and students when administered frequently (Brookhart, 2009; Chappuis, Stiggins, Chappuis, & Arter, 2012; Popham, 2008), and it should assist the teacher in determining instructional adjustments with the purpose of improving student achievement (Brookhart, 2009; Davies, 2009; Fisher & Frey, 2007; Willhoft, 2012). Not only should formative assessments occur on a regular basis, they should also vary in presentation each time they are administered (Popham, 2008; Sadler, 2010), and they do not need to be graded or weighed heavily (Brookhart, 2009; Fisher & Frey, 2007; Matthews, 2007; Popham, 2008; Sadler, 2010). Formative assessments provide the stage for proper placement within standards and course objectives (Doubet, 2012), and they give teachers immediate results to discern the most current issues with student understanding (Bartlett, 2012). Educators would do well to remember assessment is a process—not the final product (Matthews, 2007). The goals of formative assessment are to improve scores, understanding, confidence, quality learning for students, and to guide instruction (Edwards, 2013; McMunn, Schenck, & McColskey, 2003; Doubet, 2012; Stiggins, 2007), and huge strides are gained when students are properly measured (Stiggins, 2006).

There are various forms of assessment, and as long as they maintain focus on the primary goals, they will produce the desired results (Chappuis, Chappuis, & Stiggins, 2009). Differentiated instruction is possible when many types of assessments are put into regular practice. Student-made tests are a great way for teachers to understand how students think and learn in order to adequately meet the students' needs (Lutz, 2011). Quizzes can be set up as formative measures, not graded or intended to make or break a student, but rather act as a barometer for both the student and the teacher to assess

knowledge and progress toward achievement (Fleenor, Lamb, Anton, & Donen, 2011; Matthews, 2007). Teachers who have used a check-off system for homework as formative assessment rather than a recorded letter grade have found it helpful (Fisher & Frey, 2007; Greenstein, 2012; Villa, Thousand, & Nevin, 2008; Wormeli, 2006). When assessments are part of the grading and educational process, they are more beneficial to students than the traditional once-and-for-all final exam at the end of a course (Stiggins, 2007). According to Chappuis, Chappuis, and Stiggins (2009), "the five keys to assessment are:

- Clear purpose
- Clear learning targets
- Sound assessment design
- Effective communication of results
- Student involvement in the assessment process" (p. 14).

There is a strong correlation between coaching athletes and teaching students regarding the philosophy and practical application of regular assessments (Chappuis, 2009; Fleenor, Lamb, Anton, & Donen, 2011; Wormeli, 2011). Coaches conduct practices to introduce and train players in the fundamental skills of a sport, providing ample time for them to learn and apply the lessons prior to a game. Similarly, teachers continually help students identify their skill levels and provide them with tools and strategies for improving their learning and skill before a big test or assessment (Chappuis, 2009; Sadler, 2010; Willis, 2006).

Just as a coach cannot remove a player in the middle of a big game to teach a fundamental skill, a teacher cannot expect to instruct a student in the basics during a test.

Such instruction should be provided during class time. Teachers should give valuable and applicable feedback to the students throughout the course of instruction, which are assessments for learning (formative assessments) before giving high-stakes tests (summative assessments), or assessments of learning (Chappuis, 2009; Fisher & Frey, 2007). Summative assessments provide students the opportunity to demonstrate their knowledge (Chappuis, Stiggins, Chappuis, & Arter, 2012) and all they have learned in preparation for the big game (Davies, 2009; Fleenor, Lamb, Anton, & Donen, 2011; Rom, 2011). A good test should address the concepts taught and be weighed according to the degree of difficulty and time spent on the material. Another sensible strategy is to not overburden students with multiple tests in a specific time segment (Salend, 2011). Formative assessments increase academic achievement (Schmoker, 2011) and, unlike summative assessments, are given with plenty of time for students and teachers to make changes to improve knowledge and understanding (Varlas, 2012). Both kinds of assessment are needed for the preparation and analysis of student knowledge and to adequately meet the educational needs of students (Chappuis, 2009).

Regular and effective feedback propels students into learning and promotes a positive mindset for the rest of the course (Chappuis, 2009; Dweck, 2006; Patron & Smith, 2011; Ramey, VandeVusse, & Gosline, 2007; Vatterott, 2011). To support and increase student learning, a system of immediate and specific feedback is preferred (DiBattista, Gosse, Sinnige-Egger, Candale, & Sargeson, 2009; Marzano, 2008; Tovani, 2012; Wormeli, 2012). This creates an opportunity for educators to quickly and accurately identify students' needs and address them accordingly (Marzano, 2008;

Wagner, 2002). Too often, a teacher's only feedback to students is the recorded grades for all homework assignments.

A more productive grading practice is to give feedback through students' homework using other strategies to assess and build upon the level of the students' mastery without assigning grades (Matthews, 2007; Schimmer, 2012; Wormeli, 2006). The art of understanding feedback does not come naturally to students. Feedback should come to students in both written and oral forms from their teachers for optimal results (Brookhart, 2008). The use of technology can also help teachers give feedback more quickly and more easily. Using Google Docs to teach students to write a thesis is an example of how technology increases teacher productivity and makes feedback more accessible to students (Gullen & Zimmerman, 2013). It is important for the teacher to know how to use and model proper feedback methods. For example, a teacher can provide an open forum where anonymous, incorrect answers are shared and discussed orally (Brookhart, 2008; Wormeli, 2012). Providing quality feedback on regular basis is essential to enhancing the educational process, increasing student engagement and understanding, and improving summative scores (Chappuis, 2009; Dweck, 2006; Patron & Smith, 2011; Ramey, VandeVusse, & Gosline, 2007; Vatterott, 2011). Good and timely feedback of a student's knowledge and understanding is a much more valuable educational tool than a final number or letter grade (Tomlinson, 2011).

There are several preferred alternatives to traditional evaluative methods. One alternative is allowing students to take some control of their education and empowering them in their own learning by teaching them how to reflect on and present personal assessment information (Fisher & Frey, 2007; Matthews, 2007). A teacher allows

students to monitor their own progress and collect artifacts to be placed in personal portfolios (Chappuis, 2009; Matthews, 2007; Popham, 2011). Students may have the ability to choose how many assignments to fulfill and submit based on a teacher's recommendation and standards. It promotes an intrinsic drive for education and accomplishment when students are able to take responsibility for their own learning (Martin, 2013).

Student motivation increases through the process of self-examination and tracking personal progress (Brookhart, 2009; Wormeli, 2012). Regular reflection enhances a student's ability to engage in deeper discussions and assess their own learning (Fisher & Frey, 2007; Popham, 2011). These skills can be practiced via settings, such as student-led conferences. When students have a vested interest and take personal responsibility to present and explain their own progress, the result is a high level of motivation and accountability (Brown & Knowles, 2007; Chappuis, 2009; Kinney, 2012). Another source for learning new grading techniques is computer software. Some software programs allow a teacher to mark only the grades, while others give options allowing teachers flexibility in recording procedures, such as breaking down a unit into quizzes (formative assessment), homework, and midterms (Gullen, Gullen, & Erickson-Guy, 2012; Marzano, 2006). Some programs are easier to use than others and may provide teachers the ability to manipulate their grading procedures within the software. An effective grading software system will go beyond the mere spreadsheet calculation of averages, allowing teachers to record both formative and summative assessments as they relate to standards (Marzano, 2006).

Technology has also made grading more visible and accessible, and it has ensured transparency and accountability that at one time did not exist. It compels educators to be open and willing to make the best use of grades (Gullen, Gullen, & Erickson-Guy, 2012; Schimmer, 2012). Beyond choices of software and the use of technology, Danielson (2002) said policies and practices must be in place in a school and district to encourage students to succeed. The perception of students, parents, teachers, and administration should be that hard work is superior and more valuable than natural endowment or luck (Brookhart, 2009).

Assessment and grading practices should provide students and parents the clarity pertaining to specific levels of achievement and performance, while at the same time promoting the intrinsic motivation to learn (Popham, 2011). Involving students in the processes of assessment, record keeping, goal setting, and communication is important to the students' intellectual and personal growth (Brookhart, 2009; Fisher & Frey, 2007; Van de Walle, 2001). When students become involved in the assessment processes, the results are increased buy-in, motivation, and effort (Brookhart, 2009; Fisher & Frey, 2007; Martin, 2013; Schimmer, 2012; Schurr, 2012). It also prepares them to succeed in higher education (Khan & Slavitt, 2013). Effective assessments provide enough evidence to evaluate and measure student learning (Bailey & Jakicic, 2012; Moss & Brookhart, 2009; Schimmer, 2012; Willis, 2006). Record keeping helps monitor the progress of student learning. Goal setting presents a target toward which students strive. Communication about assessment offers parents, teachers, and students an essential connection to discuss matters of education (Marzano, 2007; Matthews, 2007; O'Connor & Wormeli, 2011; Stiggins, Arter, Chappuis, & Chappuis, 2006). Allowing students the

opportunity to monitor their own academic achievements makes them feel more involved and important in the process (Khan & Slavitt, 2013). Assigning more homework does not guarantee better grades. Involving students in the assessment process, creating opportunities for peer observations, and implementing a peer assessment model with proper training enhance the educational experience and result in better grades (Rom, 2011; Schimmer, 2012). Grades, however, are not the only motivational element. The act of learning itself should become a much stronger motivator for students (Guskey, 2011; Popham, 2008; Resh, 2009; Wormeli, 2012). There are many who believe grades do not necessarily motivate all students (Kohn, 2011) and say educators should look for different ways to motivate students besides the grade (Dyrness & Dyrness, 2008; Guskey, 2011; Resh, 2009; Winger, 2005). Students need to participate in the assessment process and even help decide the measures for judgment, as this presents an invaluable experience for learning and augments the importance of a good grade (Rieg, 2007; Schimmer, 2012).

Effective alternative methods are operational in school districts as proof there are advantages to changing outdated grading practices. Implementing strategies promoting progress in learning is very successful (Erickson, 2001; O'Connor, 2007). If administrators are going to ask teachers to abandon the practice of giving zeros, then they need to provide a viable alternative, which is insisting upon work completion by the students (Guskey, 2004). Instead of the conventional zero, an "I" can be used to signify incomplete work or insufficient evidence of mastery. Students who have an "I" are required to make up the work, so they can prove an understanding and grasp of the material (Reeves, 2004). Teachers can also allow students to retake assessments in order

to acquire a better grade and improve students' understanding of the concepts. This has demonstrated a noticeable difference in both low-achieving and high-achieving students (Dueck, 2011; Gordon & Fay, 2010; Wormeli, 2011).

Contract grading as opposed to traditional grading is used to motivate students to work on assignments and learn in the process. Most submitted work does not receive a letter grade but a check signifying the assignment was completed and given to the teacher (Potts, 2010). Developing course competencies is paramount and must include intervention plans when students fail them. Competencies may include traditional tests (multiple choice and short answers), oral presentations, projects, performances, and writings (Fisher, Frey, & Pumpian, 2011). Numerous constructive options for strategies improving the academic experience for students establish the need and value for thinking outside of the proverbial box (Erickson, 2001; O'Connor, 2007).

Common assessment policies and practices within a school or throughout a district are necessary for the success of any educational strategy (Goodwin & Kirsten, 2012; Schmoker, 1999). Guskey and Jung (2012) said,

> Grading reform is a necessary piece of the move toward a standards-based orientation to education. Although numerous decisions must be made when revising report cards and it may seem daunting, the four components of grading reform are the most important first steps.
>
> - Be clear about the purpose.
> - Use multiple grades.
> - Change procedures for selecting the class valedictorian and eliminate class rank.

- Give honest, accurate, and meaningful grades (p. 28).

There has been a positive shift in the perception of grading and assessment in many school districts across the country. Many have included the use of standards and rubrics (St. Maurice & Yudchitz, 2003). Rubrics can be applied to students' performance, assignments, projects, creativity levels, and portfolios as alternative measurement tools (Brookhart, 2013; Doganay & Pinar, 2010; Schimmer, 2012). Some school districts, such as Minnetonka, Minnesota, have already transformed their grading practices. Students are allowed to retake tests, I's are given instead of zeros, and students are held accountable to finish the work. These changes have not created more work for the staff; it has been described as different work (Erickson, 2001; Patron & Smith, 2011). Teachers in rural Nebraska are incorporating formative and summative measures for grading, while, fewer of them use zeros (Stephens, 2010). Kentucky has implemented a standards-based report card for elementary and secondary schools (Guskey, Swan, & Jung, 2011). Students in pass–fail classes have reported more learning opportunities through assessments than students in multistep grading programs (Dahlgren, Fejes, Abrandt-Dahlgren, & Trowald, 2009). Letter grades in some districts have been converted to numbers on a numerical rating scale similar to one used for computing grade-point average (Randall & Engelhard, 2009). Teachers within a PLC participate in collective inquiry and experimentation to implement best practices in grading (Lawrence, 2011; Schimmer, 2012). In some higher performing schools, teachers have described an open door policy to allow others in the building the freedom to observe their grading practices (Marzano, 2008; Wilcox & Angelis, 2012). The creative and practical options for new-and-improved grading

practices are countless and provide schools with tools for achieving a greater degree of student success within the educational process (Erickson, 2001; O'Connor, 2007).

Purposeful Community

The primary goal of a middle level principal as an instructional leader is to establish a purposeful community (Waters & Cameron, 2007). The intentions of a purposeful community include encouraging students and staff to value and promote middle level philosophies and best practices for grading and assessment and advocating for the unique needs of those in this phase of education (NMSA, 2010). A purposeful community believes each staff member is capable of making a perceptible difference in the lives of students using the tools available, such as grading practices, curriculum, instructional strategies, and assessment (DuFour, DuFour, Eaker, & Karhanek, 2010). Purposeful communities also operate in a manner where the staff performs daily routines, such as grading and assessment practices that are fair and consistent and have been previously agreed upon through a collaborative process and professional development (Erkens & Twadell, 2012). This may mean the establishment of a school-wide grading and assessment policy. Cooperation in the staff environment is an essential element within a purposeful community and can be achieved when an administrator casts a vision everyone can support and implements strategies and professional development with common goals and objectives (DuFour, DuFour, Eaker, & Karhanek, 2010). For instance, there would be the high expectation for every member of a purposeful community to participate in professional development opportunities pertaining to best practices in grading and assessment and eventually to establish a uniform grading and assessment policy (Reeves, 2009). A purposeful community is personal and envelops people with a

family-like atmosphere where members know and understand one another, and it is within this context, the goal for a school-wide grading and assessment policy can be introduced, discussed, and implemented (Eaker & Keating, 2012).

In order to obtain such a goal, there are four virtues driving the framework, including affirmation, communication, culture, and ideals and beliefs. These virtues are tied to specific rules and practices leading to a purposeful community (Waters & Cameron, 2007). The goal of a purposeful community may be achieved when the virtues have been well defined and collectively followed (DuFour, DuFour, Eaker, & Karhanek, 2010). For virtuous characteristics to flourish in a school, they must be singled out and rewarded. Success breeds success and will eventually lead to a purposeful community, along with keeping focus on the goal of agreed-upon practices in grading and assessment (Waters & Cameron, 2007).

The first defined virtue is affirmation, and there are certain responsibilities tied to affirmation within a middle school principal's position. Recognizing and acknowledging failures and celebrating accomplishments when efforts are made toward standardizing grading and assessment practices are crucial (DuFour & Marzano, 2011; Waters & Cameron, 2007). Staff members will feel valued and supported when they perceive their opinions are heard and understood. They will be more inclined to try new strategies when there is the general sentiment that all ideas are on the table and appreciated and there is value attached to every individual (Eaker & Keating, 2012).

The second virtue is communication. The ability to establish strong lines of communication is vital in order to effectively exchange information. A practice associated with a middle level principal's duties to establish good communication is

assuring accessibility (DuFour & Marzano, 2011; Waters & Cameron, 2007). A leader must be open and approachable for teachers, students, and parents. A leader should also develop an effective means for teachers to be able to openly and constructively communicate with one another (Kinney & Robinson, 2005). Purposeful communities are in continuous connection with each other through various avenues of communication, which makes isolation uncomfortable. It is the only way healthy discussion pertaining to grading and assessment practices can take place, and a principal should maintain open lines of communication with staff members at all times. Open dialogue guarantees the introduction of the best options and ensures collaboration, so the final choice of policy is a mutual agreement (DuFour & Marzano, 2011).

Establishing a healthy culture is the third virtue on which a middle school principal must focus (Waters & Cameron, 2007). Promoting a healthy culture requires leadership that exudes a good sense of community and cooperation and effectively cultivates shared beliefs (Jackson & Davis, 2000). Fostering the virtue of culture consists of several associated practices. A principal should promote cooperation, have a sense of well-being, improve cohesion among staff, and relate to the community. One way for these to take place is through the establishment of family and community partnerships initiated by the school (Epstein, 2005). It is important to develop an understanding of the overall goal of the school so to encourage a shared vision of what the school could and should look like. If teachers, students, parents, community members, and administration are all working together to improve and enhance the culture of the school, the purpose of the community will be strengthened (AMLE, 2010). Getting all stakeholders on board with the goal of a uniform grading and assessment policy would make the transition of

implementation much easier (Swaim, 2005). The previous establishment of a healthy culture would ensure stakeholders support and trust a school policy decision, rather than criticize or disparage it. Utilizing key practices would also help provide the necessary guidance to achieve the goal of the school (AMLE, 2010).

Establishing strong ideals and beliefs about schooling is essential within the framework for administrative leadership (Waters & Cameron, 2007). A leader must be able to operate a school and impart the ideals and beliefs in such a way where they are understood, embraced, and followed. The practices associated with ideals and beliefs reflect the overall operation of the school (Waters & Cameron, 2007). Middle level principals should hold strong professional beliefs about the school, the teaching, the grading and assessment practices, and the student learning characterizing their school (Fullan, 2003). These beliefs must be shared with teachers, students, and parents in order to strengthen the school's goals (Swaim, 2005). Demonstrating these practices through consistent behavior, collaboration, and having the courage to make difficult decisions helps stakeholders embrace the ideals and beliefs (Fullan, 2003; Kinney & Robinson, 2005; Kouzes & Posner, 2006). Professional development is also a powerful tool for a principal to convey ideals and beliefs about grading and assessment practices. If administrative leaders model the ideals and beliefs through the means of best practices, the purposeful community can be reached (DuFour, 2004; Eaker & Keating, 2012).

A purposeful community develops in time because of good leadership (Waters & Cameron, 2007). With increasing expectations in middle level education along with decreasing budgets, the pressure on schools has magnified, and the middle school principal's role has become more complex. Principals set goals and display the courage to

attain the skills necessary to lead and manage students and teachers through the middle level grades (Eaker & Keating, 2012; Fullan, 2003; Kinney & Robinson, 2005; Waters & Cameron, 2007). Relational trust, which places a strong emphasis on the regard for each staff member, is acquired through time and will eventually help to propel the goal to another level (Fullen, 2003). Maintaining open dialogue with other administrators and establishing mentorship strata are valuable practices able to also accelerate a school's timeline for reaching its goals (DuFour & Marzano, 2011). Having attainable goals supported by strong virtues and guided practices provides hope for the future and a chance for students to flourish and find success in education through a purposeful community (Fullan, 2003).

Professional development is the ideal venue to introduce new goals, foster a collaborative climate, and challenge the status quo (Waters & Cameron, 2007). Confronting faulty grading and assessment practices and presenting better options are most effectively accomplished through professional development (Brookhart, 2011; Erickson, 2011; McDaniel, 2010). One of the duties of a middle school principal is to lead in making necessary educational changes (Marzano, Waters, & McNulty, 2005; Waters & Cameron, 2007), and professional development is the tool to assist those processes. In order to sustain change in any method or policy, a staff must participate and benefit from frequently offered, quality professional development (Allen, 2005; Brookhart, 2011; Dyb, 2012; Erickson, 2011; Fisher, Frey, & Pumpian, 2011).

Professional development also promotes collaboration and should be entrenched into a staff's regular schedule through teaming and PLCs (DuFour, 2004; Eaker & Keating, 2012). This can be regularly scheduled through early release time on a particular

day of the week, prep-hour meetings, or rotation of regular departmental meetings. Collaborative administrators have the ability to share knowledge with their staffs because they have committed themselves to further learning and opportunities to practice methods and tools for education (Eaker & Keating, 2012; Marzano & Waters, 2009). Professional development programs are best created and executed by a cooperative team of administration and teachers (DuFour & Marzano, 2011; Nesin, 2005). When the administrative leadership provides a setting to identify best practices in grading and assessment and allows a team approach to investigate and determine next steps, it inherently promotes a staff's capacity to collaborate, hence the formation of a purposeful community (Wilcox & Angelis, 2012).

Professional Development

Professional development is essential to transform ineffective and erroneous grading practices (McDaniel, 2010). It is the most effective means for influencing change in grading practices (Brookhart, 2011; Erickson, 2011), and one of the duties of a middle level principal is to lead in making necessary educational changes (Marzano, Waters, & McNulty, 2005; Waters & Cameron, 2007).

Nearly all increases in student achievement can be connected in some way to high-quality professional development (Barr & Parrett, 2007). Continuous professional development leads to better classroom practices and success (Teague & Anfara, 2012) and guarantees grades become a true representation of student knowledge and achievement (Aronson, 2008; Moss & Brookhart, 2009). Training models promoting meaningful conversations about the nature of student achievement, systematic assessment practices, and accurate and equitable grading methods have a positive and lasting impact

on teachers and their students (Brookhart, 2011; Erickson, 2011; Roorda, 2008; Stiggins, 2007).

Variations in teachers' grading practices are due to the lack of formal training and professional development (Guskey, 2009; Schimmer, 2012), and only a small number of educators know and understand the numerous options in grading practices and the ramifications on student learning (Guskey & Jung, 2012; Stiggins, Arter, Chappuis, & Chappuis, 2006). Professional development is the necessary element for introducing and implementing innovative, research-based educational strategies. It is also the primary way to ensure their maintenance and continued implementation (Dyb, 2012). Professional development is the most powerful and effectual tool for implementing and sustaining the needed change in grading and assessment practices in classrooms, schools, and districts across the country (Dyb, 2012; Roorda, 2008).

Collaboration is crucial to the successful implementation of any strategy because a joint effort provides the greatest opportunity for research to discover optimum practices and proven results (Matthews, 2007; Schimmer, 2012). Collective decision making within PLCs should be a lasting commitment (Proulx, Spencer-May, & Westerberg, 2012). Schools reporting high scores and achievement focus on collaboration, not control, and practice sharing duties rather than tenure; it is not a one-person job (Riddle, 2012). Problem solving, examining new ideas, attempting different practices, and experiential learning are indicators of a collaborative environment (DuFour, DuFour, Eaker, & Karhanek, 2004; DuFour, DuFour, Eaker, & Many, 2010; Wilcox & Angelis, 2012). Opportunities for small-group interaction, experimentation, and reflection ensure optimal and long-lasting results. Studies and experience show new concepts and practices will fail

despite their worthiness, if the chance to discuss and practice them within a collaborative group is not provided (Stiggins, Arter, Chappuis, & Chappuis, 2006; Barr & Parrett, 2007). Further intensive training through collaboration is an ideal option to identify future school leaders (Wilcox & Angelis, 2012). Feelings of loneliness and isolation in the career of education are easily remedied when teachers collaborate within PLCs and participate in professional development (DuFour, 2004; DuFour, DuFour, Eaker, & Many, 2010; Schimmer, 2012). Schools and teachers who are collaborative and committed to continual professional learning and development keep the goal of student achievement central. They demonstrate a willingness to learn and apply the best methods in education, which inevitably leads to greater understanding of the students they teach (Barr & Parrett, 2007; Schimmer, 2012).

Professional development should offer a recurrent model not restricted to conferences, seminars, workshops, books, and classroom observations (Stiggins, Arter, Chappuis, & Chappuis, 2006). A school staff should collaborate and contemplate the goals and purposes of grading methods and the real and practical effects in daily class operations (Fisher, Frey, & Pumpian, 2011). Professional development and collaboration should be entrenched into a staff's regular schedule through teaming and PLCs (DuFour, 2004; Eaker & Keating, 2012). Collaborative administrators have the ability to share knowledge with their staffs because they have committed themselves to further learning and opportunities to practice methods and tools for education (Eaker & Keating, 2012; Marzano & Waters, 2009). Administrative leadership and professional development are among the contributing factors of capacity building (equipping and investing in future leaders) and collaboration (working together) at the middle level of education (Barr &

Parrett, 2007; Wilcox & Angelis, 2012). Professional development programs are best created and executed by a cooperative team of administration and teachers. When administrative leadership provides a venue to identify best practices in grading and assessment and allows a team approach to investigate and determine next steps, it inherently promotes a staff's capacity to collaborate (Wilcox & Angelis, 2012).

Professional development may be customized to the unique needs and intentions of the specific organization (Aseltine, Faryniarz, & Rigazio-DiGilio, 2006; Zmuda, Kuklis, & Kline, 2004) and should meet individual requirements and personalities (Marzano, Norford, Paynter, Pickering, & Gaddy, 2001). It facilitates the capability for broaching the conversation about changes by setting a constructive stage for progression (Reeves, 2011). Presenters could tackle a wide spectrum of topics, ranging from student achievement, successful assessment tools, and accurate grading practices to social amendments (Roorda, 2008). It is beneficial for those who have taken classes or attended conferences to present their findings to their staff as a form of professional development. Teaching the material solidifies concepts for the presenters and simultaneously synchronizes the staff and provides the necessary training (Wormeli, 2012). School administrators may consider reaching out to measurement specialists to play a leading role in teacher development (Aronson, 2008). Professional development may look and sound different for any given group but will still offer invaluable results if integrated into the routine or schedule of a program (Aseltine, Faryniarz, & Rigazio-DiGilio, 2006; Zmuda, Kuklis, & Kline, 2004).

Continual opportunities for professional development and collaboration are vital to the introduction and maintenance of effective grading and assessment practices (Dyb,

2012; Roorda, 2008). Frequently offered professional development keeps topics in the minds and conversations of teachers instead of becoming tabled and forgotten (Wormeli, 2012). The best professional development is ongoing and prioritizes student comprehension, job issues, and the needs of the school population as a whole and of its individuals (McMunn, Schenck, & McColskey, 2003). In order to sustain change in any grading and assessment practices, a staff must participate and benefit from frequently offered professional development (Allen, 2005; Brookhart, 2011; Dyb, 2012; Erickson, 2011; Fisher, Frey, & Pumpian, 2011). Grade-level groups, interdisciplinary teams, and subject area departments are stakeholders of the professional development infrastructure to ensure quality maintenance within a school (Teague & Anfara, 2012). Professional development should be part of a school's philosophy and thus incorporated into daily operations, including opportunities to practice desired skills (Moore, 2012; Proulx, Spencer-May, & Westerberg, 2012). Learning to collaborate, organize, implement and sustain change, motivate, gather and evaluate data, lead meetings, and handle conflict is most effectively taught and maintained via frequent and continual professional development (Moore, 2012).

When districts are ready to revise grading and assessment practices, quality professional development must be provided for the schools' staffs to attain the desired results (McMunn, Schenck, & McColskey, 2003). Quality, successful professional development must have the following elements:

- Uniform and consistent
- Recurrent
- Emphasis on long-range

- Builds knowledge base in others
- Tackles whole school and individual needs
- Collaborative
- Applicable and comprehensible
- Conducted on campus
- Student learning is paramount
- Based on research
- Professional development in curriculum and instruction
- Buy-in from all stakeholders
- Values constructive input (Elmore, 2002; McMunn, Schenck, & McColskey, 2003).

Conclusion

It is the responsibility of those in the teaching profession to continue to learn and stay apprised of current research and data to give students the best possible education (Schimmer, 2012). New times call for new methods, and education is no exception; educators must embrace the wave of the future rather than tenaciously hold onto the past (Robinson & Aronica, 2009). As the transition to a standards-based education becomes the reality, leaders must implement the necessary changes in grading practices in order to stay up to date (Guskey & Jung, 2012). Waters and Cameron (2007) identified focus as a leadership responsibility for principals, which helps to establish clear goals for the school year and to keep those goals a priority for the school. Highly effective schools and leaders offer regular opportunities for participation and instruction through professional development (Teague & Anfara, 2012). Faculty buy-in and support are also crucial to the

success of any consequential changes, and teachers will make the necessary efforts to implement change when there is support from the school district and professional development activities (DuFour & Marzano, 2011; Fisher, Frey, & Pumpian, 2011; McMunn, Schenck, & McColskey, 2003; Proulx, Spencer-May, & Westerberg, 2012). Training produces success, and staff development has been proven to increase teacher knowledge, skills, and confidence in crucial matters of grading and assessment (Matthews, 2007). Success in the implementation of uniform grading practices is achievable when there is staff support and consistent, ongoing, quality professional development (Gullen, Gullen, & Erickson-Guy, 2012). Waters and Cameron (2007) explained second-order change as an apparent break from past practices, requiring a new skill or knowledge set to be acquired that may not align closely with an individual's personal belief and value system. Change offers fresh perspective, the chance to try innovative ideas, a time to grow both personally and professionally and to become more efficient in teaching (Schimmer, 2012). Professional development offers the best chance for the successful implementation, maintenance, and sustainability of changes in the fundamental approaches to grading and assessment practices in education, forever impacting the lives of our students (Dyb, 2012; Gullen, Gullen, & Erickson-Guy, 2012; McDaniel, 2010).

Chapter III

Design and Methodology

The goal of this study was to evaluate the perceptions of grading and assessment practices at FRMS and determine the impact professional development opportunities had on changing those perceptions. Educational leaders, administrators, and teachers are faced with questions regarding the best ways to motivate students and accurately report their progress (Popham, 2011). Change is inevitable and necessary at times, offering fresh perspective and new ideas, and it is the responsibility of those in the teaching profession to continue to learn and stay apprised of current research to give students the best possible education (Schimmer, 2012). In-service training was provided for the staff members of FRMS, with the hope that they would recognize the need for a change in past grading practices and understand how to select and apply the tools necessary for moving forward with sound grading practices. The professional development training challenged the staff to consider the advantages of adopting a uniform method for grading and assessment in which grade reports would reflect achievement on standards, while establishing a separate process for reporting student behavior, engagement, and work ethic (Brookhart, 2011; O'Connor, 2007; Yesbeck, 2011).

Research Design

The research strategy selected was a quantitative study. The quantitative study was conducted at FRMS, located in a rural community in the northwestern part of the United States. FRMS teachers had inquired about and explored the possibility of creating consistency in grading and assessment practices. There were many grading and

assessment philosophies represented by teachers at FRMS. Professional development was implemented to allow teachers to investigate best practices for grading and assessment. The initial professional development opportunity for FRMS was an in-service training, including five expert teachers who presented sessions pertaining to grading and assessment practices. FRMS teachers attended these classes as part of their district's regularly scheduled professional development and were encouraged to ask as many questions as necessary for clarity throughout the sessions.

The first presenter introduced the topic of grading summative assessments (particularly tests) based on a proficiency scale. The presenter explained that in order to be fair to students, the summative assessment should be given only after there has been adequate preparation. The assessment should be holistically based on standards, and a reassessment should be given when necessary. A typical test in the presenter's classroom was organized and designed based on standards, and within each standard, the questions were written at all the levels of cognitive sophistication (below basic, basic, proficient, and advanced). Language provided by the new Bloom's Taxonomy (see Appendix C) or The Depth of Knowledge (see Appendix D) was used to structure the questions and guarantee a hierarchy of proficiency levels. The presenter displayed a grade book (PowerSchool) and demonstrated how to convert proficiency levels into percentages (see Appendix E). The presenter shared a self-assessment (see Appendix F) students typically answered prior to taking the test to determine where they landed on the proficiency scale. Later, they used the self-assessment as a study tool if they scored a basic or below basic understanding. The presenter said, at times, there was not enough evidence to determine proficiency or there was an incomplete status on an assignment, which acted as a

placeholder until the assessment was finished and submitted. The presenter concluded by sharing example tests (see Appendix G) and methods for conducting reviews and retakes.

The second presenter discussed student responsibility regarding grading and assessment in three stages. The first topic addressed was, “Assessment for Learning: Where Am I Going?” The presenter explained in this model, the students took a section prequiz and evaluated their understanding of the concepts based on their answers, using the Reviewing My Math Results check-off list (see Appendix H). It was placed with the prequiz inside their Math notebook. The next section was, “Where Am I Now?” During this stage, the students were encouraged to closely examine their prequiz review analysis and determine whether they understood the material, whether they made minute errors, or whether they clearly did not understand the material. This provided students the opportunity to see where they needed to pay special attention. It individualized the learning for each student and created a self-evaluation of the standards covered on the test. Students used notebooks to quickly and easily reflect on what they had covered and where they may have required additional help. The final stage presented was, “How Do I Bridge the Gap?” Based on individual proficiency for each standard, students played a game or participated in a specific task to strengthen a particular concept (see Appendix I). If students had no missing assignments and had completed a review of the standard(s) they missed and reviewed with the teacher, they were given an opportunity to retake a test or redo an assignment. This became their final score; there was not another chance offered. The presenter concluded by saying the use of assessments for learning within a grading and assessment policy guaranteed students were taking responsibility for their learning and success. Students also had multiple exposures to the standards (see

Appendix J) and learning targets (see Appendix K), which provided them with more opportunities to gain and demonstrate proficiency on those standards.

The third presenter spoke about formative assessments versus summative assessments. Formative assessments included notes, warm-ups (see Appendix L), homework, class work, quizzes, worksheets, and group-based projects. Summative assessments encompassed a variety of tests and projects. The presenter detailed a grading scale (see Appendix M) and said formative assessments should consist of 20% or less of the students' grades, while the remaining 80% should be based on summative evidence of the students' knowledge of state standards. Tests are graded according to proficiency levels on a scale of one to four: one is below basic–novice level, two is basic–apprentice level, three is proficient–practitioner, and four is advanced–master of grade level or higher standards. Student progress monitoring and reflection of assessment and learning are also extremely important. The presenter offered several examples of how students recorded and monitored their levels of achievement on each of the learning strands (see Appendix N). The presenter explained instead of assigning zeros or extra credit, there were many alternatives provided to increase student comprehension of standards. Additional practice and reteaching prepared students to retake tests and become proficient, raising their achievement levels and, in turn, their class grade.

The fourth presenter's session related to two very broad areas of grading and assessment: challenging conventional practices and separating factors within a grade. It began with a short video (*Implementing Standards-Based Grading*, by Aaron Moiso, copyright 2010), contrasting usual grading practices with new-and-improved grading practices. It emphasized the mathematical unfairness of zeros in the grade book and

addressed the idea of treating nonworking students as a behavior issue rather than a grading issue. The presenter discussed the difference between practice (for student knowledge) and assessment (of student knowledge) and explained the problem of using extra credit to boost a student's grade without actually improving a student's knowledge base. The presenter described the mental and physical blocks involved in changing a school's grading and assessment policies and the common concerns teachers have when confronted with the idea of no longer giving zeros, no longer grading every assignment, and providing multiple opportunities to demonstrate performance on assessment. The presenter detailed a program piloted at the presenter's school that implemented lunch detention, after-school detention, and suspension consequences for chronic nonworkers (see Appendix O). The presenter stressed the importance of providing regular feedback to students rather than simply assigning a letter grade and the necessity for making any grading policy school-wide in order to ensure changes are successful and long lasting. The presentation ended by offering an example of a school-wide grading policy (see Appendix P).

The fifth and final presenter addressed the need for teachers to create rubrics for each assessment, so they are able to focus only on the information they are trying to assess. This is especially important in the English and Language Arts classroom because it is too easy for Language Arts teachers to get caught up in the mechanics of English when they should be looking for something more complex.

Three types of rubrics were presented:

- Single-Skill Rubric (see Appendix Q): allows the teacher to be very focused in the assessment and to assess one standard thoroughly.

- Multiple-Skill Rubric (see Appendix R): is for different types of writing, such as research writing requiring a student to exhibit multiple skills (i.e., comprehension of a text, quoting and paraphrasing, interpreting–inferring information gleaned from a text, and citing information). This entails a bit more work in order to create and to grade, but it also provides the students a thorough set of criteria without having to tell them what to write in each paragraph.
- Kid-Friendly Rubric (see Appendix S): offers an option for translating standards-based language into kid-friendly terms, which clarifies exactly what the students are being asked to accomplish on a given piece of writing.

The presenter discussed the realities of the considerable amounts of time and intensity required, in addition to the expectations and professional obligations necessary for valuable grading. If thorough work is expected, thorough grading is necessary. The presenter explained rubrics enhance efficiency and the ability to prioritize what is being assessed to focus directly on students' achievement and comprehension of the material.

The administrator at FRMS encouraged teachers to take baby steps as they began to reform their assessment practices. The administration assumed a hands-off approach, hoping teachers would take risks without fear of formal evaluations and observations. During the nine weeks following the professional development intervention, the principal of FRMS conducted five additional grading and assessment training sessions for the staff with lessons developed and provided by the researcher (see Appendices T, U, V, W, and X). The meetings were attended every other Wednesday and each session lasted 45 minutes. Attendance was mandatory, due to the scheduled early release time set aside by the district for professional development. Prior to each meeting, teachers were assigned

outside readings (see Appendices T, U, V, W, and X). Three team leaders plus administrators rotated as facilitators for each of the five meetings. This model was highly beneficial for the teachers because they were learning directly from their peers and supervisors (Copland & Knapp, 2006). The format for the trainings was a corporate gathering for the initial delivery of information. Teachers then dispersed according to grade level teams for further discussion (see Appendices T, U, V, W, and X). FRMS has one team per grade level for sixth, seventh, and eighth grades. Specialist teachers, such as special education teachers, were assigned to one of the grade level teams. The FRMS principal reported positive responses to the material (see Appendix AA).

Participants

The study of best practices of grading and assessment was implemented at FRMS located in a rural community in the northwestern part of the United States, with an enrollment of 335 students in three grade levels, consisting of sixth through eighth graders. The demographics of the school were 53% qualified for school honor rolls, 11% earned straight A's, 96.2% attended daily, 23% were limited English, 72% were on free or reduced lunch, 33% qualified for food stamps, 15% were determined mobile, 8% received special education services, 15% attended after-school tutoring, and 32% were enrolled in enrichment classes for Reading and/or Math. The school discipline included 0% students expelled; 0% students attended school court; 0% students had alcohol, tobacco, or drug violations; 36 days of out-of-school suspensions were served; 42 days of in-school suspensions were served; 0% students served time at the detention center over the past two years; and five students were on probation. Demographics for FRMS parents included 93% attended parent–teacher conferences, 16% received e-mail progress

reports, 54% attended back-to-school night, 52% earned a high school diploma, and 17% earned a postsecondary degree. The teaching staff consisted of 22 teachers, including six males and 16 females. The average years of teaching experience among participants were 15.8 years, with three teachers holding master's degrees, and one was a national board-certified teacher (see Appendix Y).

All teachers at FRMS participated in the in-service training for grading and assessment within PLCs as part of their professional development module. The collaborative climate of FRMS was conducive to participation in a PLC setting. The researcher conducted the presurvey (see Appendix A) and postsurvey (see Appendix B) during the fall semester of 2012.

Data Collection

The data collection process began by administering two surveys to teachers at FRMS. These surveys were completed through an on-site professional staff development course on grading and assessment, and all teachers were active participants. Prior to the distribution of the surveys to the FRMS staff, the researcher conducted a pilot study with 12 teachers from three middle schools in the northwestern part of the United States. Based on the pilot study, a few questions were revised and three questions were discarded.

Data was collected to evaluate teachers' perceptions and beliefs about grading and assessment. The researcher collected presurvey data (see Appendix A) prior to an introductory meeting in the fall semester of 2012. The introductory meeting in the fall was in-service style, during which presenters taught participants about best practices in grading and assessment. A second postsurvey (see Appendix B) was distributed near the

end of the semester in the fall of 2012. Both surveys utilized the software Qualtrics, an online survey and data collection program. These surveys were analyzed quantitatively. The presurvey asked the staff about their perceptions regarding grading and assessment. The surveys also asked if the staff would support a school-wide grading and assessment policy in the future.

The response options in the survey used a 5 point Likert scale: strongly agree, agree, neutral, disagree, and strongly disagree. The Likert Scale was used to allow the individual to express how much they agreed or disagreed with a particular statement. These ordinal scales measured attitudes of agreement or disagreement of individuals, asking them to respond to a series of questions on certain topics. The Likert Scale is the most popular method for recording survey research answers and is not the same as the Likert-type scale, or rating scale. In the Likert-type scale, questions are referred to as Likert items, while the Likert scale is the sum of many of the items (McLeod, 2008; Tanner, 2012).

Teachers had the option of not completing the survey. A reasonable time was allotted to complete each survey. The researcher obtained data from the surveys through the password protected software program.

Analytical Methods

The study of best practices in grading and assessment was conducted at FRMS, located in a rural community in the northwestern part of the United States. FRMS had an enrollment of 335 students in three grade levels, consisting of sixth through eighth grades. As a requirement for this course, the researcher gained approval from the Human Research Review Committee at Northwest Nazarene University before completing the

data collection portion of this study. Permission was also granted by the principal of FRMS, as well as the Superintendent of the Mountain West School District #222 (see Appendix Z).

During a common planning day, such as a collaboration day or early release day, participants completed a detailed survey of 23 questions (see Appendix B), seeking to determine the perceptions of multiple grading and assessment issues. Recurring themes from the survey helped to identify perceptions as the data was analyzed quantitatively. Qualtrics was used to issue the surveys, and the quantitative data was inserted into the Statistical Package for the Social Sciences (SPSS) software (Version 18) for calculation (IBM SPSS, 2013). The research question was analyzed using a Mann–Whitney U test and the functions in SPSS to perform the analysis. In SPSS, the procedure divided the participants for study by forming two grouping variables called "Presurvey Group and Postsurvey Group" and designating one group a value of "1" and the other group a value of "2" (McLeod, 2008).

The Mann–Whitney U test is one of the most well-known, nonparametric hypothesis tests used for determining whether one sample has larger values than the other. The Mann–Whitney U test is administered to evaluate differences between two independent groups when the dependent variable is either (a) ordinal or (b) interval. The Mann–Whitney U test is the nonparametric alternative to the independent *t*-test. The initial data from the Mann–Whitney U test must first be combined into a set of $N = n_a + n_b$ elements. The next step is to rank from lowest to highest, even those values tied in rank. From there the rankings are sorted into two separate samples (McLeod, 2008).

The reliability of the research at FRMS was important. Roberts (2006) explained reliability is the extent to which a certain model, method or strategy, when applied in varying conditions and/or settings, will generate comparable findings. In statistics, reliability is an essential concept determining the precision of measurements. Statistical reliability determines whether or not the experiment is reproducible (Creswell, 2008). The researcher checked for reliability on grading and assessment by certifying the questionnaires used would yield similar results if they were ever repeated in another school setting during a different time of year, and so forth. A Cronbach's Alpha (or coefficient alpha) was submitted through SPSS to ensure the reliability of the survey instrument. According to Tanner (2012), a Cronbach's Alpha is the most reliable statistical test when one test is administered at a time. Gliem and Gliem (2003) said when using a Likert scale, a Cronbach's Alpha must be used to ensure the internal consistency (i.e., responses are not random). George and Mallery (2003) provided the following scoring outcome guidelines when using a Cronbach's Alpha:

- " . . . > 0.9—Excellent
- > 0.8—Good
- > 0.7—Acceptable
- > 0.6—Questionable
- > 0.5—Poor
- < 0.5—Unacceptable" (p.231).

The validity of the research is just as important as the reliability. Creswell (2008) and Roberts (2006) described validity as the degree of accuracy by which the measurements of concepts are represented. The internal validity regarding the research of

grading and assessment proved valuable and worthy of use as external research for other future studies. This was accomplished by using questionnaires previously tested in past studies and measured valid.

Limitations

This study was not in any way intended to be an evaluative measure of the participants' performance or competence. This study was to explore the grading and assessment perceptions held by the teachers at FRMS. The teachers had limited exposure to best practices for grading and assessment. A limiting factor of this study was the location. It was conducted in one rural school with a relatively small teaching staff of 22 to 25 people. Another limitation was the presurvey and the postsurvey were conducted three months apart, not allowing a great amount of time for teachers to explore grading and assessment practices workable for them. Also, six sessions may not have provided enough information, research, or scenarios. Another limiting factor was the exclusion of a qualitative portion in the study. A missing piece was a questionnaire concerning the reasons for responses. This study used one testing instrument (the Mann–Whitney U test), when another test may have, combined with the Mann–Whitney U test, may have delivered greater revelations of the data.

Role of the Researcher

My responsibilities as the researcher were to conduct the research in a manner that would not bring harm to the participants, to the district, or to me, and to coordinate a professional development module addressing the research topic, grading, and assessment. As the researcher, I organized a one-day, in-service training for the staff of FRMS after analyzing the presurvey by using five core teachers who had extensive experience in

grading and assessment practices. During the in-service training, five grading and assessment expert teachers presented many of the best practices cited in the review of literature. I acted as a consultant (without payment) for the staff of FRMS throughout the semester, regarding the next steps in grading and assessment and provided them with six additional professional development lessons.

My role as the researcher in this study may have been different than some. Because I had taught for 10 years at FRMS, I had already built trusting relationships with some of the staff. I did not start from scratch like other researchers may have experienced. I certainly had to invest time to maintain those relationships; however, I expected an efficient working environment and timeline due to these prior relationships.

My background in education gave me some bias. I have been the principal of Lone Star Middle School since it opened its doors in August 2008. I have been in public education for the past 21 years, all of which have been in the middle school arena, and the past 13 years I have served as an administrator. I used my bachelor's degree from Evangel University (Springfield, Missouri) to teach Health and Physical Education, while pursuing my master's degree in Educational Administration from Idaho State University. After being a principal in the Nampa School District for 10 years, I completed my Educational Specialist degree from Northwest Nazarene University in May of 2011. I am currently enrolled in the Doctorate of Education in Educational Leadership program at NNU and plan to graduate in May 2013.

As a current board member for the Idaho Middle Level Association and past president (2005–2007), I seek to equip, unite, and inspire middle level educators across the state of Idaho. I received the Idaho Middle Level Educator of the Year award for

Region 5 in 2000–2001 and the Idaho Middle Level Educator of the Year for 2009–2010, was recognized as a National Association of Elementary School Principals Transformational Leader in 2010, and was presented by the Idaho Association of Secondary School Principals in conjunction with the National Association of Secondary School Principals as outstanding Idaho Secondary Principal (Middle Level) in 2010–2011.

I used my extensive educational experience to conduct the research at FRMS in a professional and ethical manner.

Conclusion

My hope is for FRMS to be able to evaluate the data collected in this study and determine the benefits of employing common practices and implementing a uniform grading and assessment policy for the school. Since the research was conducted with the FRMS staff in mind, I am hopeful they will reach a conclusion in favor of common grading and assessment practices and agree they are essential for optimal student achievement and growth. Continued professional development will help them learn to apply the tools necessary to move forward with sound grading and assessment practices and to achieve the goal of implementing a uniform school-wide grading and assessment policy. A commitment to quality professional development will ensure the success and long-term presence of any adopted policy.

Chapter IV

Results

This chapter presents the findings of the study and analysis of the data. Included in this chapter are summaries of descriptive statistics for areas related to the survey responses. The purpose of this quantitative study was to analyze the perceptions of grading and assessment practices and the impact of increased opportunities for professional development with the staff at FRMS.

The central research question for this study was: Can grading and assessment perceptions among middle school educators change through professional development?

Overview of Analysis

Twenty-two teachers (six male and 16 female) from FRMS were asked to take the voluntary presurvey and postsurvey (see Appendices A and B). Sixteen teachers completed the presurvey, and two chose not to give consent to use the data for the research. This allowed for the use of 14 teacher responses in the data for the presurvey study, and 13 responses on question 18. Nineteen teachers completed the postsurvey, and one teacher chose not to give consent to use the data for the research. This allowed for the use of 18 teacher responses in the data for the postsurvey study.

A Mann–Whitney U test was performed using SPSS (IBM SPSS, 2013) on each presurvey question and each postsurvey question. The results are displayed in Table 1. Question 1 is not included in the data because it was designated for consent only.

Of the 22 questions, four questions (questions 2, 3, 15, and 16) were determined through the Mann–Whitney U test to be statistically significant, which is $p = < .05$.

Eighteen of the questions were determined to not be statistically significant; however, several responses showed strong shifts in the data results.

Table 1

Mann–Whitney U Test

Question	Significance *P* = < .05
1	
2	**.018**
3	**.026**
4	.148
5	.068
6	.660
7	.243
8	.752
9	.193
10	.157
11	.139
12	.407
13	.620
14	.062
15	**.013**
16	**.035**
17	.983
18	.156
19	.086
20	.868
21	.242
22	.236
23	.212

A Cronbach's Alpha (or coefficient alpha) test was administered to measure reliability of the survey instrument's results (see Table 1). The rating scale established that both the presurvey and postsurvey had an excellent rating. Results measured .981 for the presurvey and .989 for the postsurvey (see Tables 3 and 4). The letter *N* = number of survey questions (items) on the presurvey and postsurvey.

Table 2

Cronbach's Alpha Rating

Cronbach's Alpha	Rating
> 0.9	Excellent
> 0.8	Acceptable
> 0.6	Questionable
> 0.5	Poor
< 0.5	Unacceptable

Table 3

Presurvey Report

Reliability Statistics	
Cronbach's Alpha	*N* of Items
.981	22

Table 4

Postsurvey Report

Reliability Statistics	
Cronbach's Alpha	*N* of Items
.989	22

Percentages were calculated and graphed (see Table 5) per question according to the Likert 5-point scale. Distributions of response percentages ranged from 0% to 85.72% on the 22-question presurvey.

Table 5

Presurvey Response Percentages

Presurvey Questions	**1 Strongly Agree**	**2 Agree**	**3 Neither Agree nor Disagree**	**4 Disagree**	**5 Strongly Disagree**	**Totals:**
1						
2	14.29%	57.14%	28.57%	0.00%	0.00%	100.00%
3	14.29%	78.57%	7.14%	0.00%	0.00%	100.00%
4	21.43%	50.00%	7.14%	21.43%	0.00%	100.00%
5	7.14%	85.72%	7.14%	0.00%	0.00%	100.00%
6	0.00%	35.71%	42.86%	21.43%	0.00%	100.00%
7	0.00%	28.57%	21.43%	50.00%	0.00%	100.00%
8	0.00%	21.43%	42.86%	21.43%	14.28%	100.00%
9	0.00%	14.28%	64.29%	21.43%	0.00%	100.00%
10	14.29%	14.29%	35.71%	35.71%	0.00%	100.00%
11	0.00%	57.15%	35.71%	7.14%	0.00%	100.00%
12	14.29%	35.71%	7.14%	42.86%	0.00%	100.00%
13	0.00%	7.14%	50.00%	28.57%	14.29%	100.00%
14	0.00%	28.57%	50.00%	21.43%	0.00%	100.00%
15	0.00%	57.15%	35.71%	7.14%	0.00%	100.00%
16	0.00%	42.86%	42.86%	14.28%	0.00%	100.00%
17	14.28%	64.29%	14.29%	7.14%	0.00%	100.00%
18	0.00%	7.69%	38.46%	38.46%	15.39%	100.00%
19	7.14%	71.43%	14.29%	7.14%	0.00%	100.00%
20	7.14%	64.29%	21.43%	7.14%	0.00%	100.00%
21	0.00%	7.14%	14.29%	64.28%	14.29%	100.00%
22	21.43%	35.71%	7.14%	28.58%	7.14%	100.00%
23	0.00%	21.43%	42.86%	35.71%	0.00%	100.00%

Percentages were calculated and graphed (see Table 6) per question according to the Likert 5-point scale. Distributions of response percentages ranged from 0% to 61.11% on the 22-question postsurvey.

Table 6

Postsurvey Response Percentages

Postsurvey Questions	**1 Strongly Agree**	**2 Agree**	**3 Neither Agree nor Disagree**	**4 Disagree**	**5 Strongly Disagree**	**Totals:**
1						
2	50.00%	44.44%	5.56%	0.00%	0.00%	100.00%
3	50.00%	50.00%	0.00%	0.00%	0.00%	100.00%
4	50.00%	27.78%	16.67%	5.55%	0.00%	100.00%
5	38.89%	55.56%	5.55%	0.00%	0.00%	100.00%
6	5.56%	33.33%	27.78%	27.78%	5.55%	100.00%
7	22.22%	22.22%	22.22%	27.78%	5.56%	100.00%
8	5.56%	27.78%	22.22%	38.89%	5.55%	100.00%
9	5.56%	11.11%	33.33%	33.33%	16.67%	100.00%
10	33.34%	22.22%	22.22%	22.22%	0.00%	100.00%
11	33.33%	38.89%	11.11%	16.67%	0.00%	100.00%
12	11.11%	16.67%	27.78%	33.33%	11.11%	100.00%
13	5.55%	11.11%	27.78%	27.78%	27.78%	100.00%
14	11.10%	55.56%	16.67%	16.67%	0.00%	100.00%
15	27.77%	61.11%	5.56%	5.56%	0.00%	100.00%
16	38.89%	27.78%	27.77%	5.56%	0.00%	100.00%
17	16.67%	61.11%	11.11%	11.11%	0.00%	100.00%
18	11.11%	11.11%	44.44%	27.78%	5.56%	100.00%
19	50.00%	27.78%	16.66%	5.56%	0.00%	100.00%
20	16.67%	44.44%	27.78%	11.11%	0.00%	100.00%
21	0.00%	11.11%	5.56%	44.44%	38.89%	100.00%
22	22.22%	11.11%	16.67%	16.67%	33.33%	100.00%
23	11.11%	11.11%	16.67%	38.89%	22.22%	100.00%

Question 1: By continuing in this survey, you give your consent to participate in this study.

The first survey question was simply for the respondents to grant permission to the researcher to use the data for analysis in the research study. The hope was to have all 22 staff members participate in this study, but there were a few who did not provide consent. Qualtrics was the survey instrument used to identify those respondents who did not give consent, and therefore, it did not calculate their responses as data in the overall results.

Table 7

Presurvey and Postsurvey Results

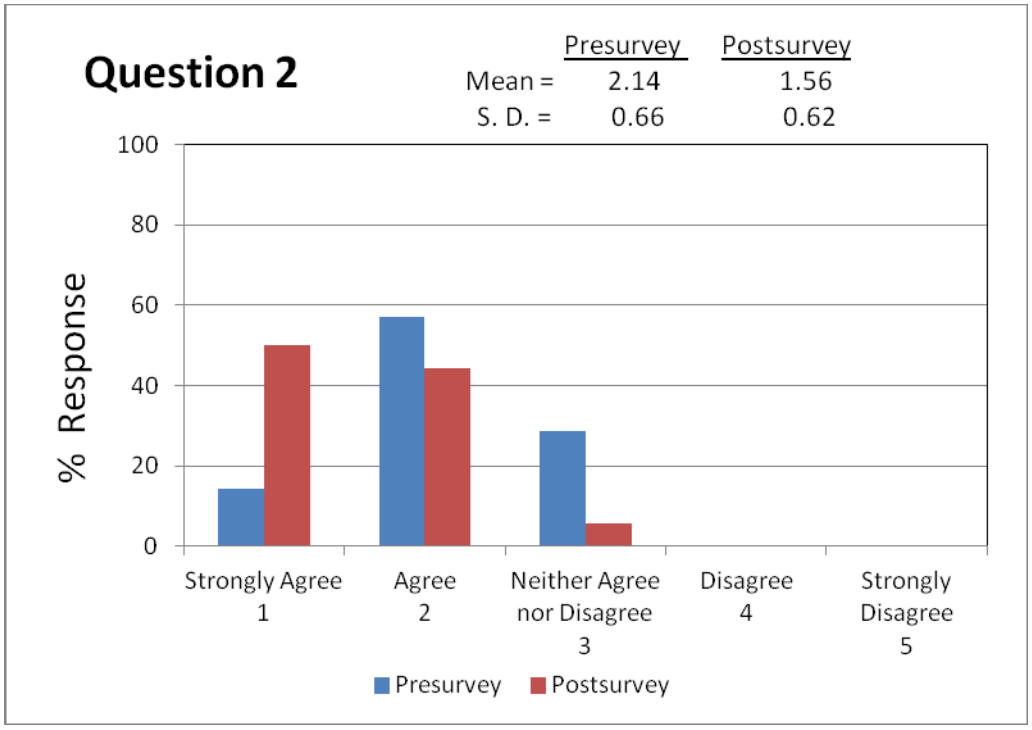
Presurvey Results (Appendix A)
Total number of respondents who completed survey: 16
Total number of respondents who gave permission to use data: 14
Postsurvey Results (Appendix B)
Total number of respondents who completed survey: 19
Total number of respondents who gave permission to use data: 18

Question 2: I support the idea of a school-wide grading and assessment policy.

As illustrated in Figure 1, the comparison of the presurvey and postsurvey results indicated a mean score shift from 2.14 to 1.56 and was statistically significant (p = .018) according to the Mann–Whitney U test. The table shows a substantial increase in the Strongly Agree category, a slight decrease in the Agree category, and a decrease in the Neither Agree nor Disagree category. There was no change in the Disagree and Strongly Disagree categories; they remained zero in both surveys. The results indicated the total number of respondents who agreed they would support the idea of a school-wide grading and assessment policy increased from 68.43% to 94.44% (total of Agrees and Strongly Agrees).

Figure 1

School-wide Grading and Assessment Policy

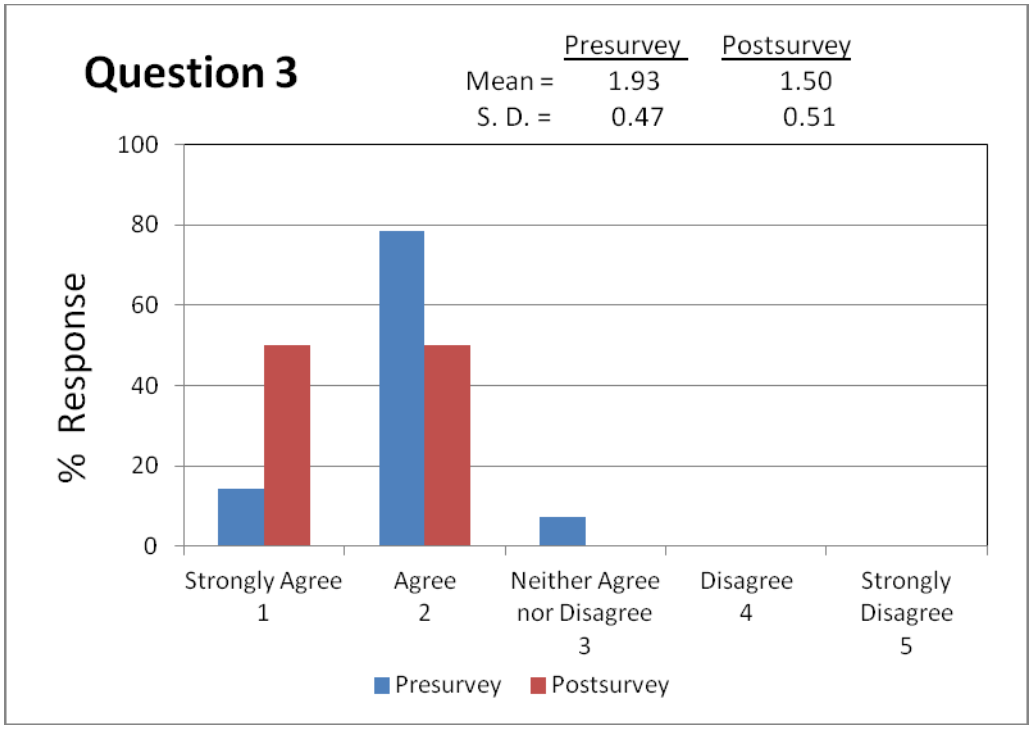


Question 3: I understand how giving zeros skews the mean (average) in calculating grades.

As illustrated in Figure 2, the comparison of the presurvey and postsurvey results indicated a mean score shift from 1.93 to 1.50 and was statistically significant ($p = .026$) according to the Mann–Whitney U test. The table shows a substantial shift from the Agrees in the presurvey to Strongly Agrees in the postsurvey. There were few to no Neither Agrees nor Disagrees, Disagrees, or Strongly Disagrees in the surveys. The results indicated perceptions shifted from a relatively high 92.86% to 100% Agrees (total sum of Agrees and Strongly Agrees) in respondents who said they understood how giving zeros skewed the mean in calculating grades.

Figure 2

Zeros Skew Averages in Calculating Grades

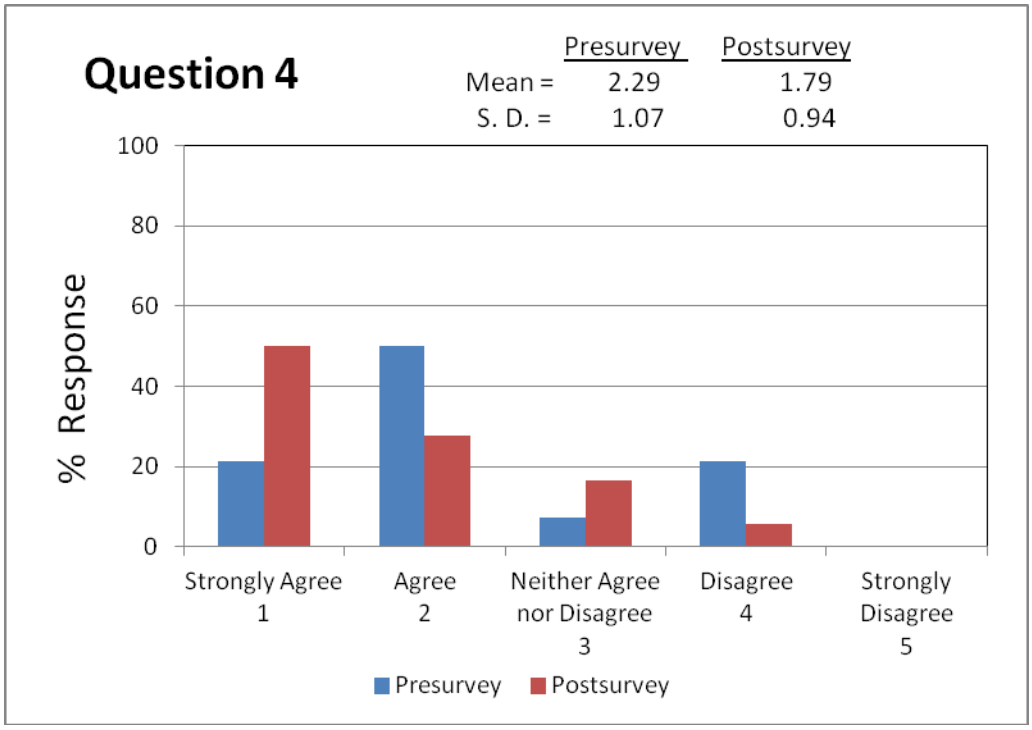


Question 4: I believe that grades should principally reflect achievement of standards.

As illustrated in Figure 3, the comparison between the presurvey and postsurvey results indicated a mean shift from 2.29 to 1.79, but it was not statistically significant according to the Mann–Whitney U test. Although there was an increase in the percent of Strongly Agrees in the postsurvey, the total of Agrees and Strongly Agrees (77.78%) was not very different from the total of Agrees and Strongly Agrees (71.43%) in the presurvey. There was a slight perceptional shift concerning grades principally reflecting achievement of standards.

Figure 3

Grades Principally Reflect Achievement of Standards

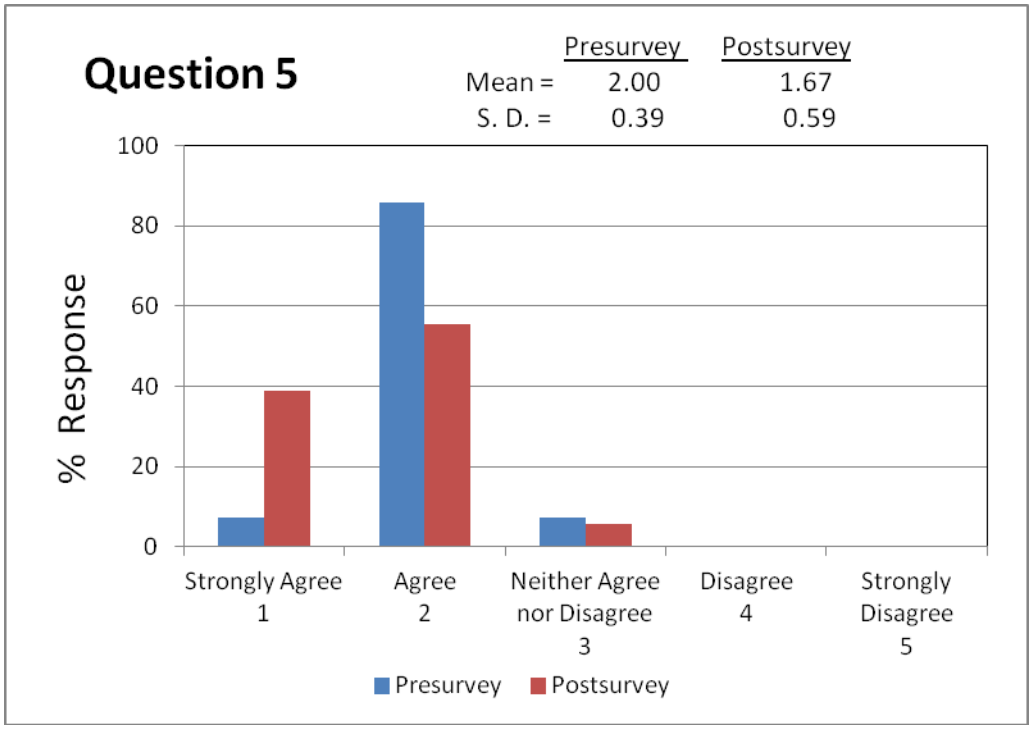


Question 5: I understand the purpose of formative assessment to be frequent feedback on progress.

As illustrated in Figure 4, the comparison between the presurvey and postsurvey results indicated a mean shift from 2.00 to 1.67, but it was not statistically significant according to the Mann–Whitney U test. Although there was an increase in the percent of Strongly Agrees in the postsurvey, the total of Agrees and Strongly Agrees (94.45%) was not very different from the total of Agrees and Strongly Agrees (92.86%) in the presurvey. There was a slight perceptional shift concerning the purpose of formative assessment as frequent feedback on progress.

Figure 4

Purpose of Formative Assessment as Frequent Feedback

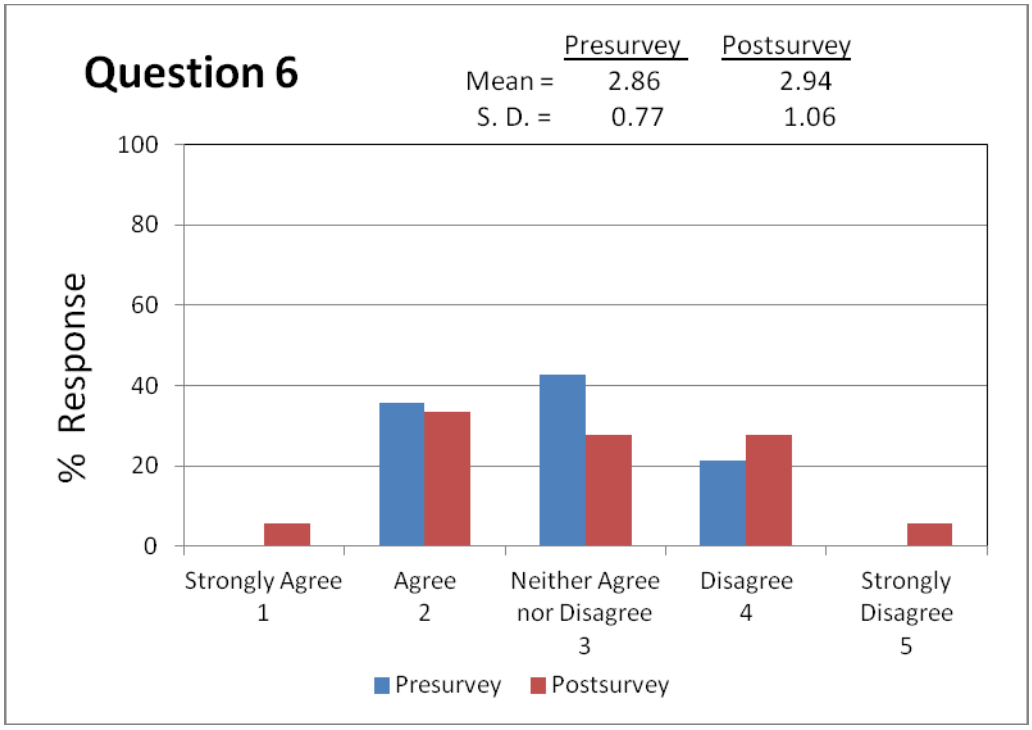


Question 6: It is objective and fair to base grades on a mean of points accumulated in a traditional scale (59% or lower = F; 60%–69% = D; 70%–79% = C; 80%–89% B; 90%–100% = A).

As illustrated in Figure 5, the comparison between the presurvey and postsurvey results indicated a mean shift from 2.86 to 2.94, but it was not statistically significant according to the Mann–Whitney U test. Although there was an increase in the percent of Strongly Disagrees in the postsurvey, the total of Disagrees and Strongly Disagrees (38.89%) was not very different from the total of Disagrees and Strongly Disagrees (35.71%) in the presurvey. There was little to no perceptional shift concerning the objectiveness and fairness of the traditional grading scale.

Figure 5

Basing Grades on a Mean in a Traditional Scale

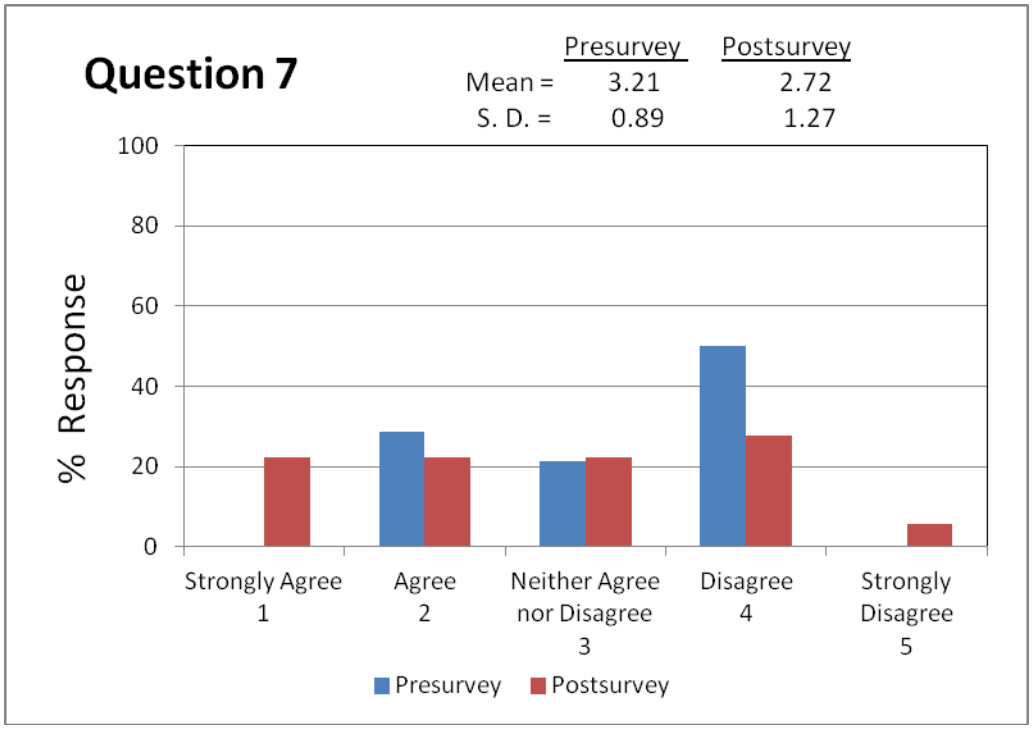


Question 7: It is okay to give a student an A if he or she has mastered a course objective but not completed all homework and class work.

As illustrated in Figure 6, the comparison between the presurvey and postsurvey results indicated a mean shift from 3.21 to 2.72, but it was not statistically significant according to the Mann–Whitney U test. There was an increase in the percent of Strongly Agrees in the postsurvey. The total of Agrees and Strongly Agrees was 44.44%, as compared to 28.57% in the presurvey, yet it did not show statistical significance on the Likert scale. There was likely a slight perceptional shift concerning the giving of an A to a student who has mastered a course objective but not completed all the course work.

Figure 6

Okay to Give an A for Mastery Without Completing Homework

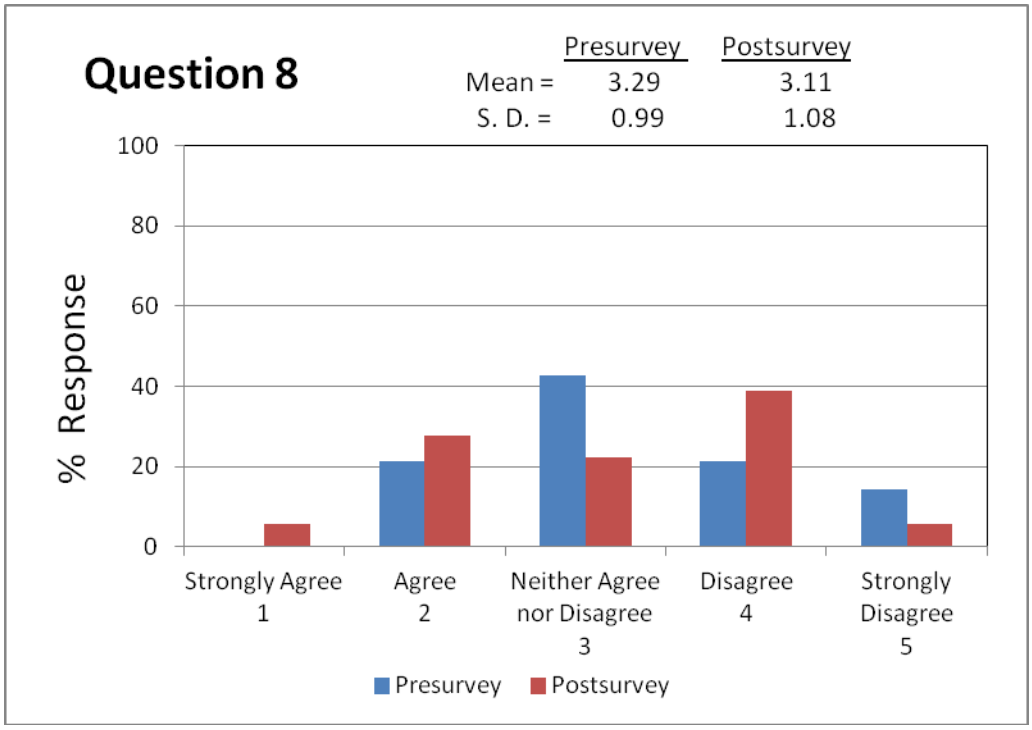


Question 8: I believe students who turn in late work should be penalized for missing deadlines.

As illustrated in Figure 7, the comparison between the presurvey and postsurvey results indicated a mean shift from 3.29 to 3.11, but it was not statistically significant according to the Mann–Whitney U test. There was an increase in the percent of Strongly Disagrees in the postsurvey. The total of Disagrees and Strongly Disagrees was 44.44%, as compared to 35.71% in the presurvey, yet it did not show statistical significance on the Likert scale. There was likely a slight perceptional shift concerning the belief in penalizing students for missing deadlines.

Figure 7

Students Penalized for Missing Deadlines

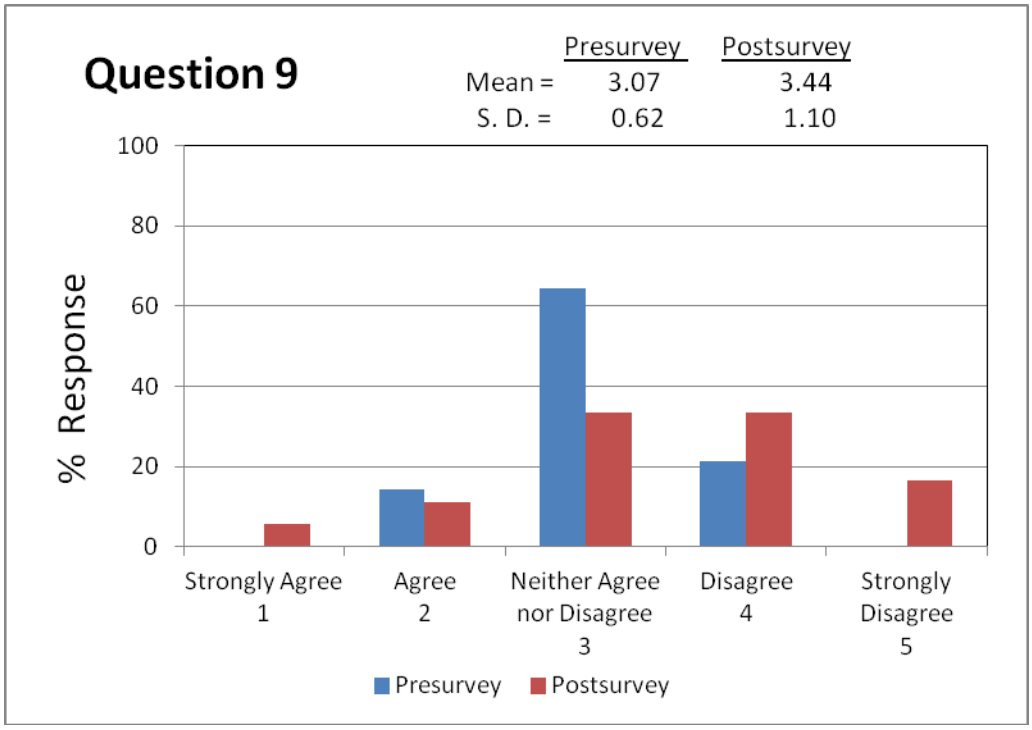


Question 9: I think teachers should be able to adopt whatever grading practices work for them.

As illustrated in Figure 8, the comparison between the presurvey and postsurvey results indicated a mean shift from 3.07 to 3.44, but it was not statistically significant according to the Mann–Whitney U test. There was an increase in the percent of Strongly Disagrees in the postsurvey. The total of Disagrees and Strongly Disagrees was 50%, as compared to 21.43% in the presurvey, yet it did not show statistical significance on the Likert scale. There was likely a positive perceptional shift concerning teachers adopting personal grading practices.

Figure 8

Teachers Should Be Able to Adopt Individual Grading Practices

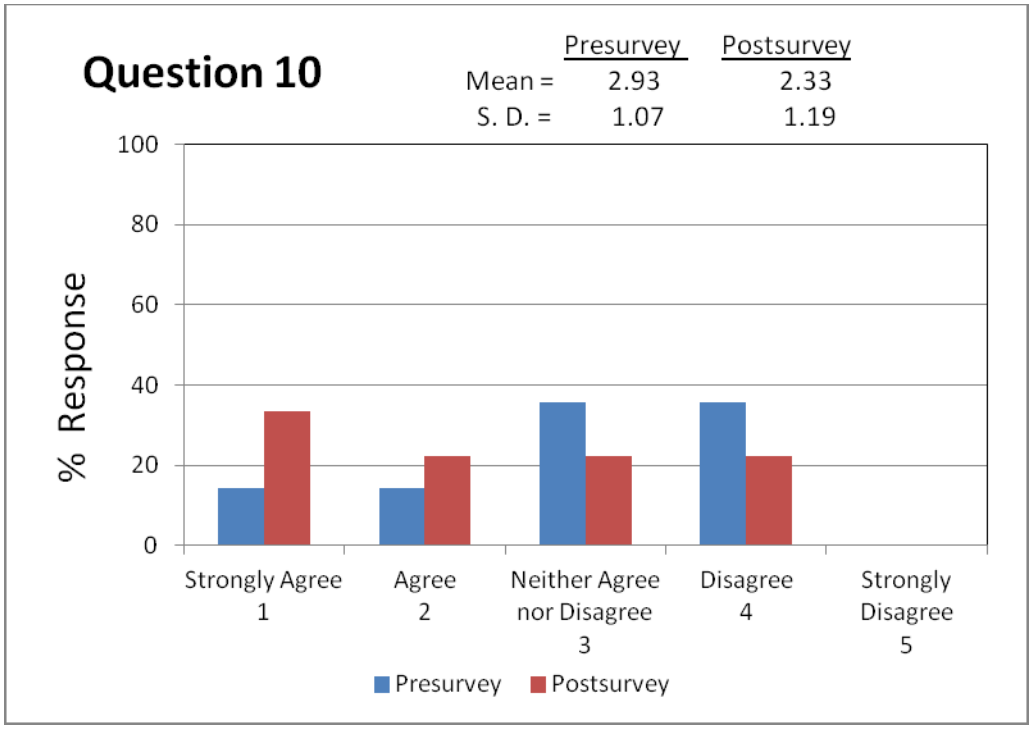


Question 10: Extra credit should not be allowed in any class.

As illustrated in Figure 9, the comparison between the presurvey and postsurvey results indicated a mean shift from 2.93 to 2.33, but it was not statistically significant according to the Mann–Whitney U test. There was an increase in the percent of Strongly Agrees in the postsurvey. The total of Agrees and Strongly Agrees was 55.56%, as compared to 28.58% in the presurvey, yet it did not show statistical significance on the Likert scale. There was likely some positive perceptional shift concerning the allowance of extra credit in classes.

Figure 9

Extra Credit Should Not Be Allowed

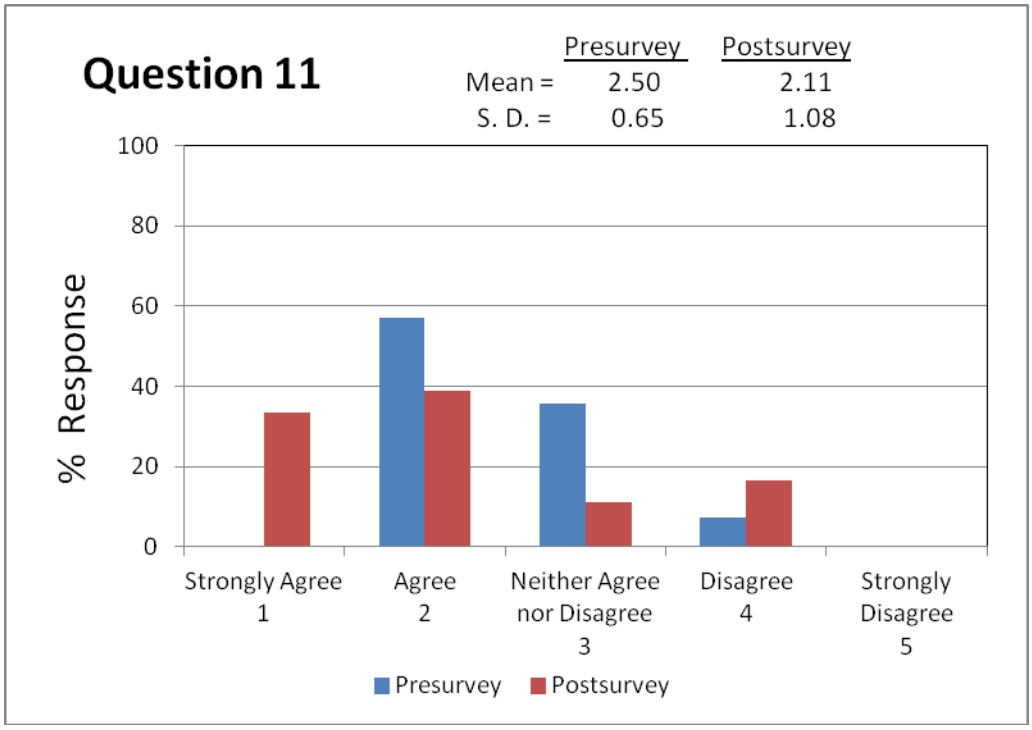


Question 11: Using a grade of incomplete is preferable to penalizing late work or giving zeros or Fs.

As illustrated in Figure 10, the comparison between the presurvey and postsurvey results indicated a mean shift from 2.50 to 2.11, but it was not statistically significant according to the Mann–Whitney U test. There was an increase in the percent of Strongly Agrees in the postsurvey. The total of Agrees and Strongly Agrees was 72.22%, as compared to 57.15% in the presurvey, yet it did not show statistical significance on the Likert scale. There was likely some perceptional shift concerning using a grade of incomplete rather than penalizing late work or giving zeros or Fs.

Figure 10

A Grade of Incomplete Is Preferable to Penalizing, Zeros, and Fs

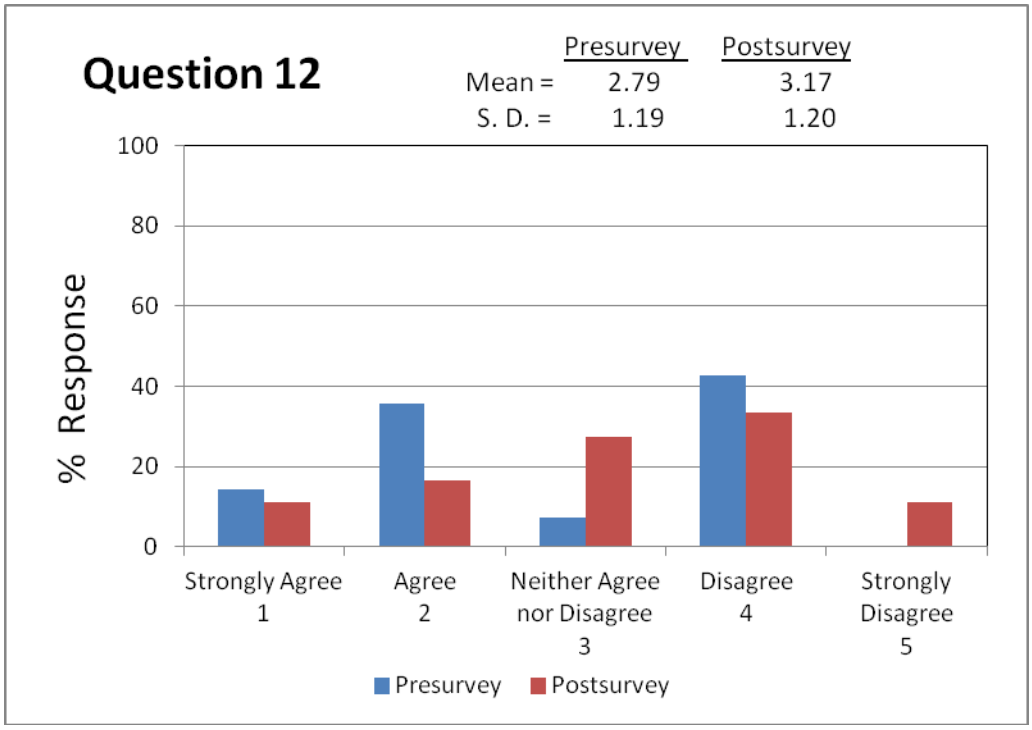


Question 12: I believe grades should reflect the influence of a student's behavior and work ethic on his achievement.

As illustrated in Figure 11, the comparison between the presurvey and postsurvey results indicated a mean shift from 2.79 to 3.17, but it was not statistically significant according to the Mann–Whitney U test. Although there was an increase in the percent of Strongly Disagrees in the postsurvey, the total of Disagrees and Strongly Disagrees (44.44%) was not very different from the total of Disagrees and Strongly Disagrees (42.86%) in the presurvey. There was not a perceptional shift concerning grades reflecting behavior and work ethic on a student's achievement.

Figure 11

Grades Reflect Influence of Behavior and Work Ethic

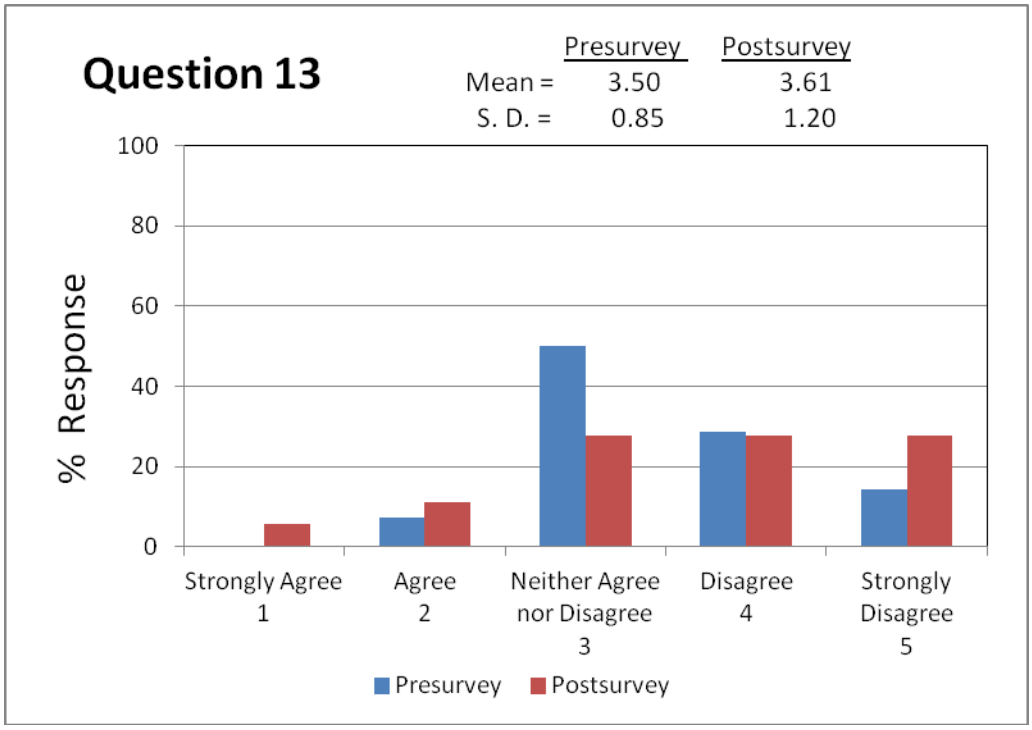


Question 13: Grades are comparative in nature and serve to sort students into ability groups.

As illustrated in Figure 12, the comparison between the presurvey and postsurvey results indicated a mean shift from 3.50 to 3.61, but it was not statistically significant according to the Mann–Whitney U test. There was an increase in the percent of Strongly Disagrees in the postsurvey. The total of Disagrees and Strongly Disagrees was 55.56%, as compared to 42.86% in the presurvey, yet it did not show statistical significance on the Likert scale. There was likely a perceptional shift concerning grades as a tool for comparing students and sorting students into ability groups.

Figure 12

Grades Are Comparative and Serve to Sort Into Ability Groups

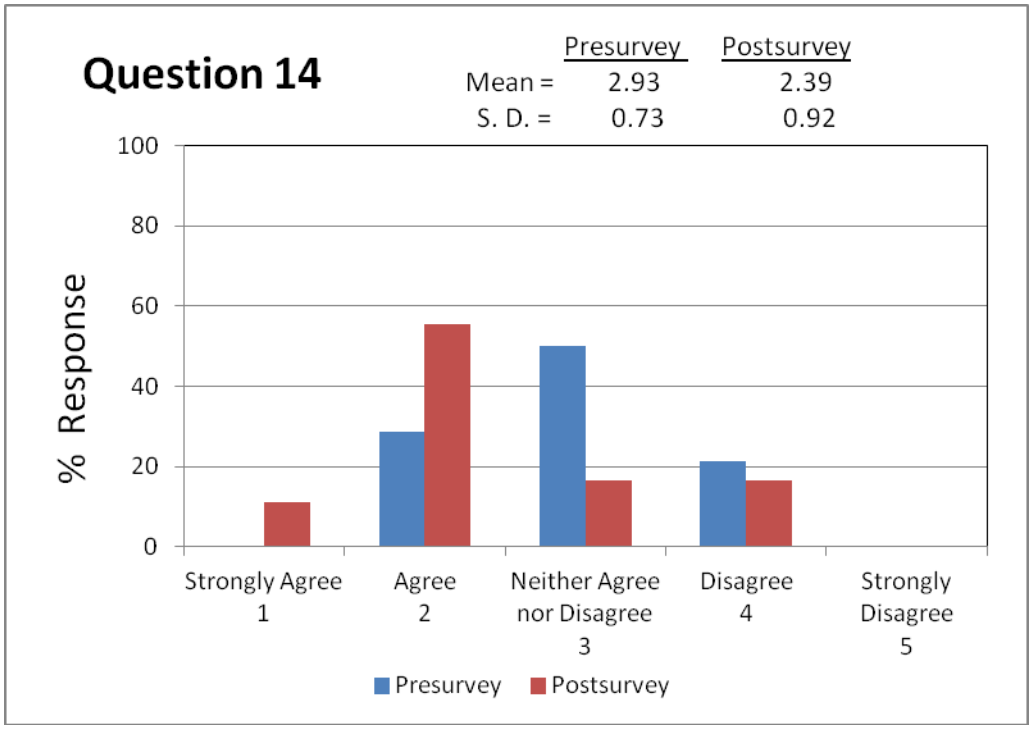


Question 14: I understand how to develop high-quality summative assessments.

As illustrated in Figure 13, the comparison between the presurvey and postsurvey results indicated a mean shift from 2.93 to 2.39, but it was not statistically significant according to the Mann–Whitney U test; although, $p = .062$ indicated the results were very close to significant. There was an increase in the percent of Strongly Agrees in the postsurvey. The total of Agrees and Strongly Agrees was 66.66%, as compared to 28.57% in the presurvey, yet it did not show statistical significance on the Likert scale. There was likely a perceptional shift concerning the understanding of how to develop high-quality summative assessments.

Figure 13

High-Quality Summative Assessments

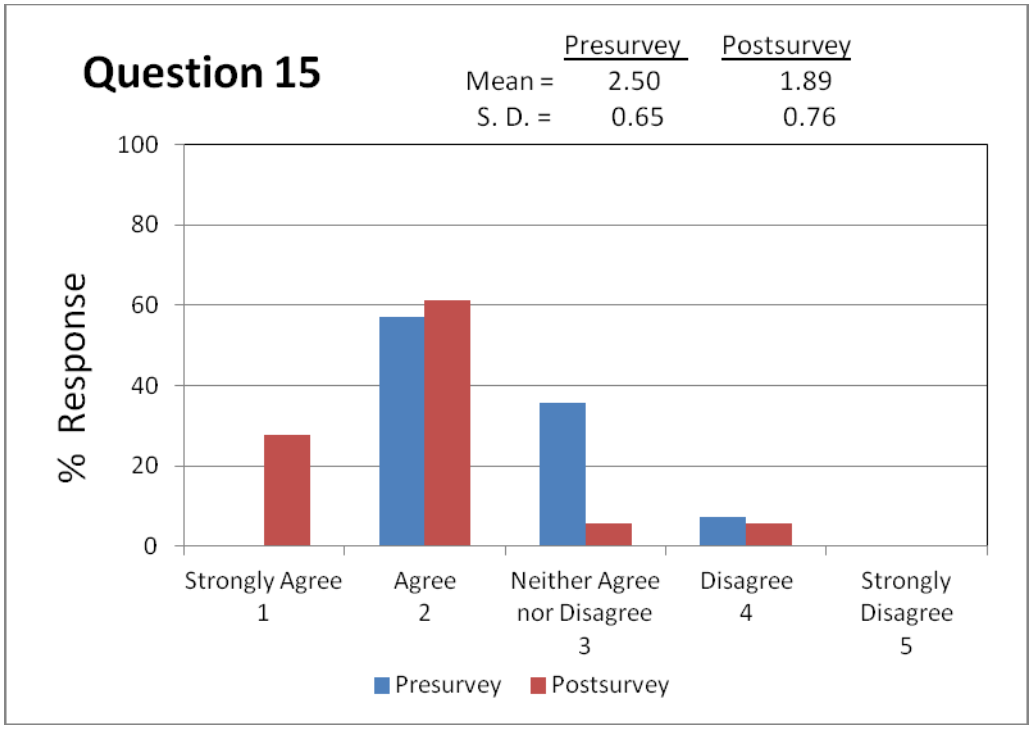


Question 15: I understand the purpose of summative assessment to be a periodic report of achievement of standards.

As illustrated in Figure 14, the comparison of the presurvey and postsurvey results indicated a mean score shift from 2.50 to 1.89 and was statistically significant (p = .013) according to the Mann–Whitney U test. The table shows a substantial increase in the Strongly Agree category, a slight increase in the Agree category, and a decrease in the Neither Agree nor Disagree category. There was a slight change in the Disagree category and no change in the Strongly Disagree category. The results indicated staff respondents had the perception that the purpose of summative assessment was to be a periodic report of the achievement of standards.

Figure 14

Summative Assessments Are Periodic Reports of Achievement of Standards

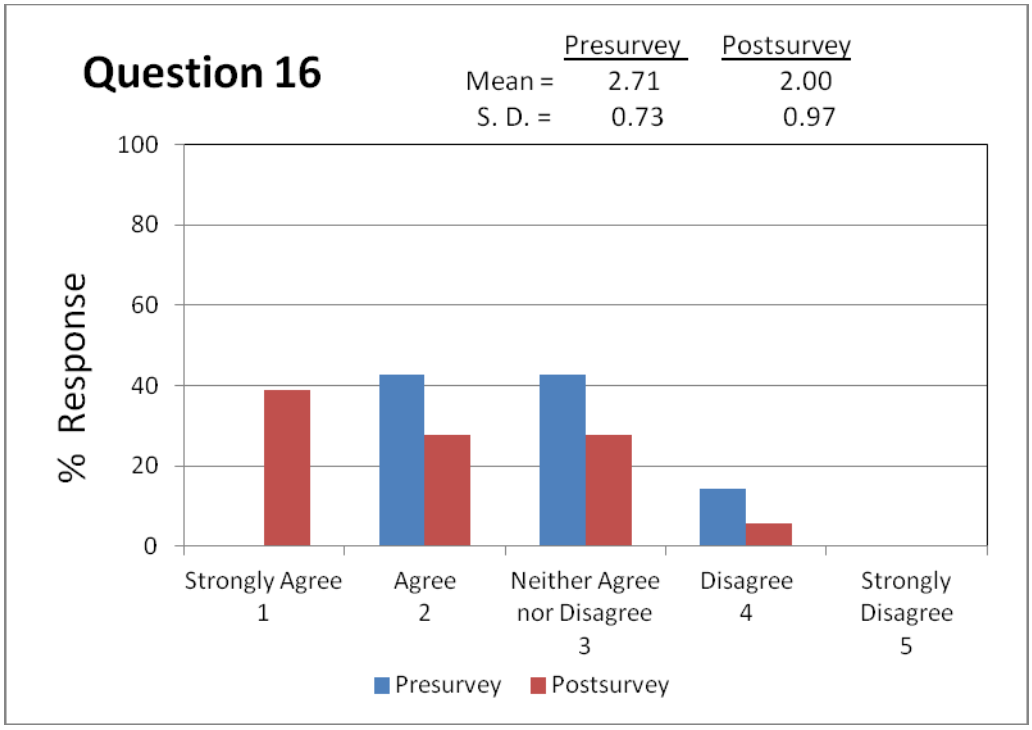


Question 16: In figuring summative grades, recent evidence is more valuable than an average of points over time.

As illustrated in Figure 15, the comparison of the presurvey and postsurvey results indicated a mean score shift from 2.71 to2.00 and was statistically significant (p = .035) according to the Mann–Whitney U test. The table shows a substantial increase in the Strongly Agree category, a decrease in the Agree and Neither Agree nor Disagree categories, and a slight decrease in the Disagree category. There was no change in the Strongly Disagree category; it remained zero in both surveys. The results indicated the staff respondents had the perception that recent evidence was more valuable than an average of points over time while figuring summative grades.

Figure 15

Recent Evidence Is More Valuable Than Average in Summative Assessments

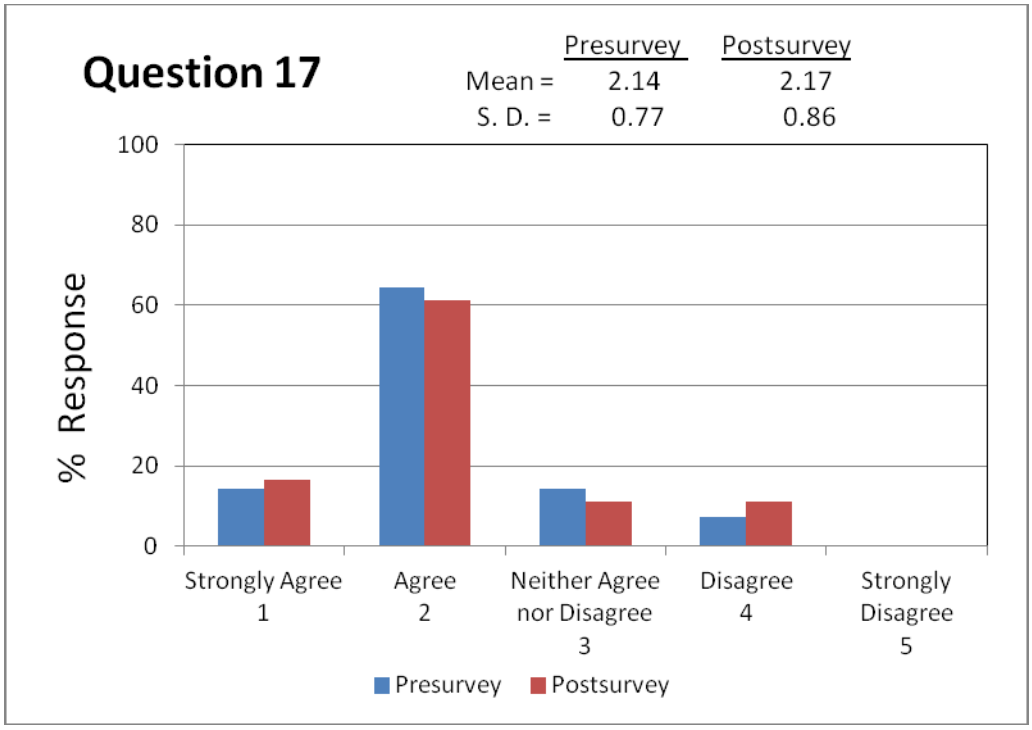


Question 17: My students are routinely involved in the formative assessment of their learning.

As illustrated in Figure 16, the comparison between the presurvey and postsurvey results indicated a mean shift from 2.14 to 2.17, but it was not statistically significant according to the Mann–Whitney U test. Although there was a decrease in the percent of Strongly Disagrees in the postsurvey, the total of Disagrees and Strongly Disagrees (77.77%) was not very different from the total of Disagrees and Strongly Disagrees (78.57%) in the presurvey. There was little to no perceptional shift of the staff respondents concerning their students being routinely involved in formative assessment of their learning.

Figure 16

Students Routinely Involved in Formative Assessment

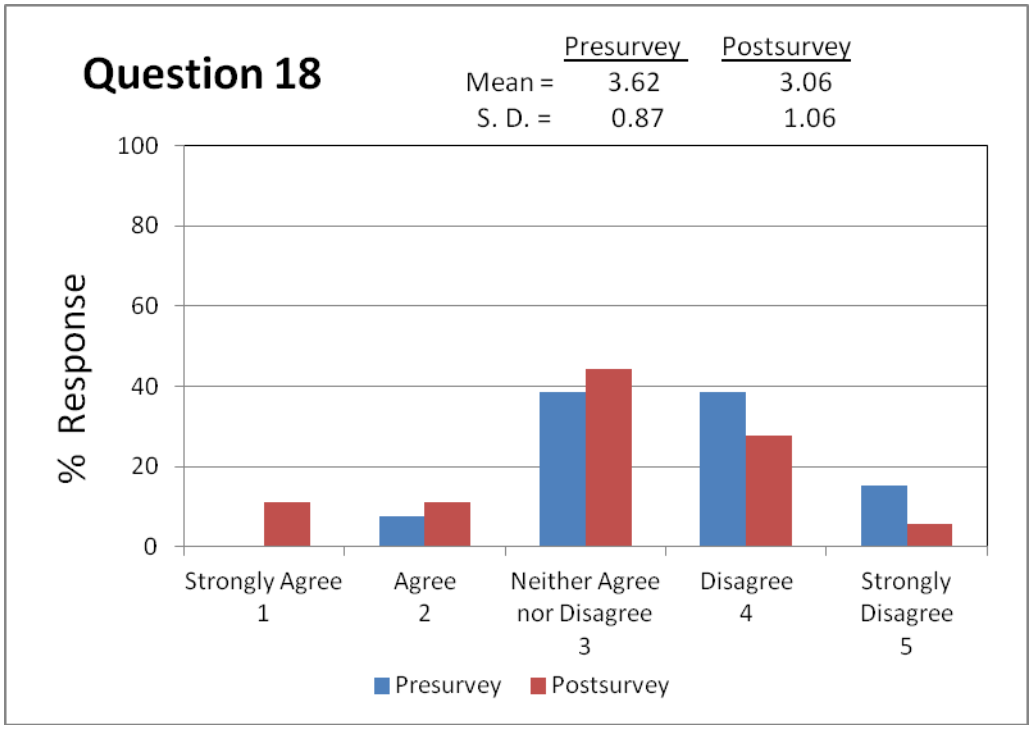


Question 18: The grading information system (e.g., PowerSchool) has interfered with my grading practices.

As illustrated in Figure 17, the comparison between the presurvey and postsurvey results indicated a mean shift from 3.62 to 3.06, but it was not statistically significant according to the Mann–Whitney U test. There was an increase in the percent of Strongly Agrees in the postsurvey. The total of Agrees and Strongly Agrees was 22.22%, as compared to 7.69% in the presurvey, yet it did not show statistical significance on the Likert scale. There was likely a slight perceptional shift with staff respondents that the grading information system (e.g., PowerSchool) interfered with grading practices.

Figure 17

Grading Information System Interfered With Grading Practices

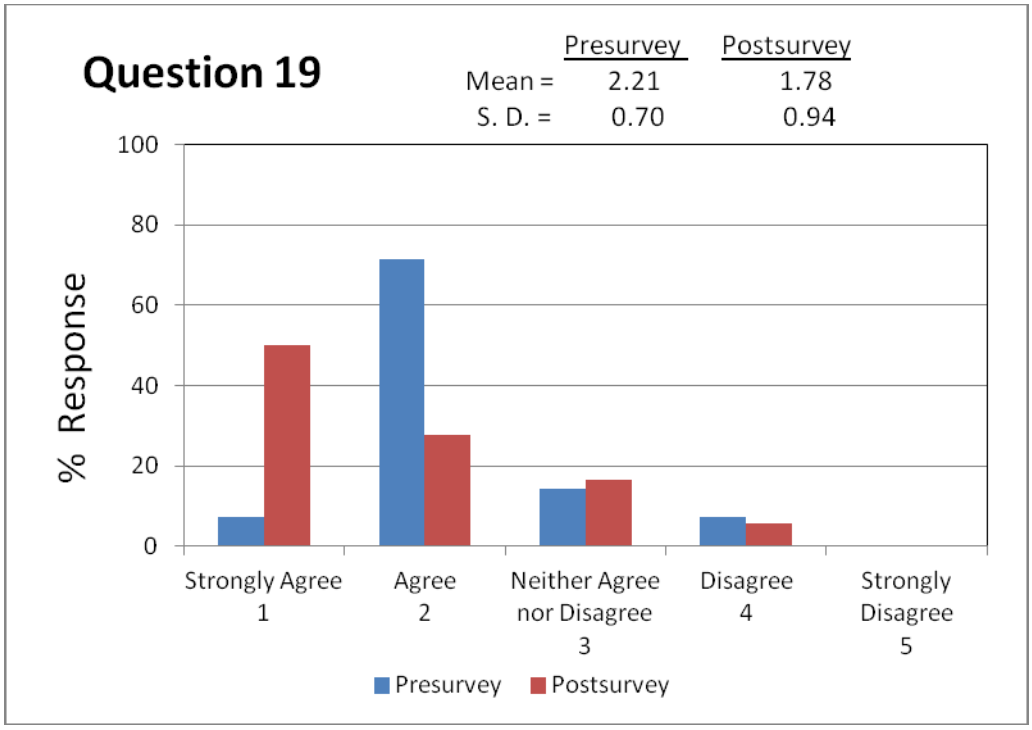


Question 19: High expectations for learning are communicated in the overall school setting.

As illustrated in Figure 18, the comparison between the presurvey and postsurvey results indicated a mean shift from 2.21 to 1.78, but it was not statistically significant according to the Mann–Whitney U test, although $p = .086$ indicated the results were very close to significant. While there was an increase in the percent of Strongly Agrees in the postsurvey, the total of Agrees and Strongly Agrees (77.78%) was not very different from the total of Agrees and Strongly Agrees (78.57%) in the presurvey. There was little to no perceptional shift concerning the communication of high expectations for learning in the overall school setting.

Figure 18

High Expectations Communicated in School Setting

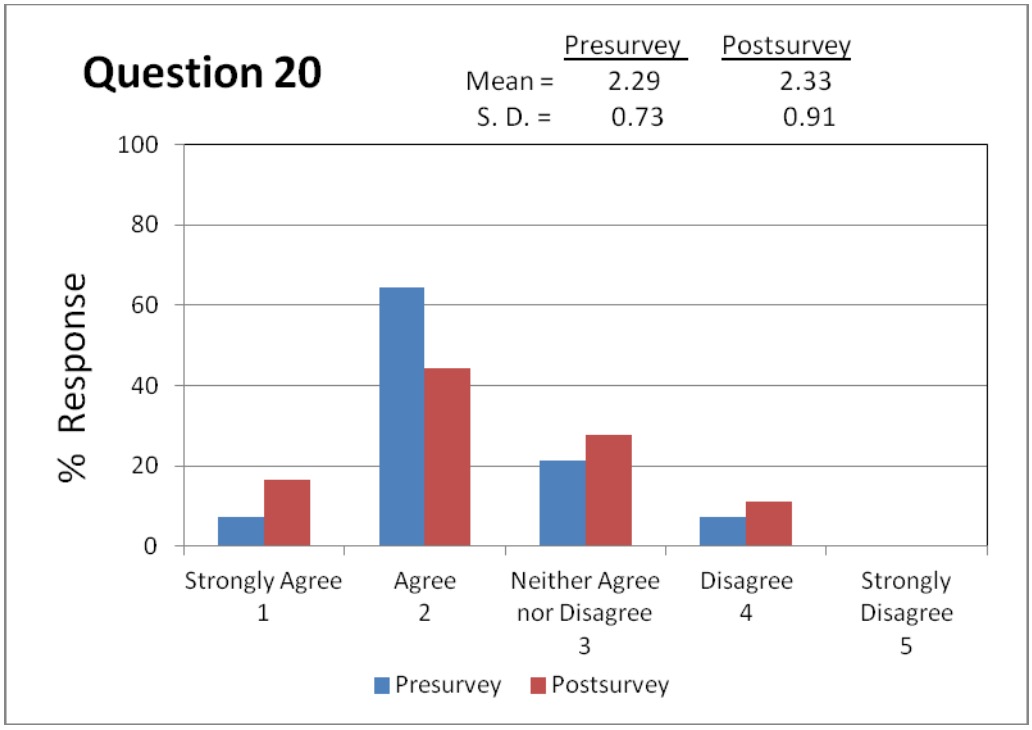


Question 20: Teachers should enter grades at least once a week into the grading information system (e.g., PowerSchool).

As illustrated in Figure 19, the comparison between the presurvey and postsurvey results indicated a mean shift from 2.29 to 2.33, but it was not statistically significant according to the Mann–Whitney U test. Although there was a decrease in the percent of Strongly Agrees in the postsurvey, the total of Agrees and Strongly Agrees (61.11%) was not very different from the total of Agrees and Strongly Agrees (71.43%) in the presurvey. There was likely a slight perceptional shift of the staff respondents regarding teachers entering grades at least once a week into the grading information system (e.g., PowerSchool).

Figure 19

Teachers Should Enter Grades Once a Week in Grading System

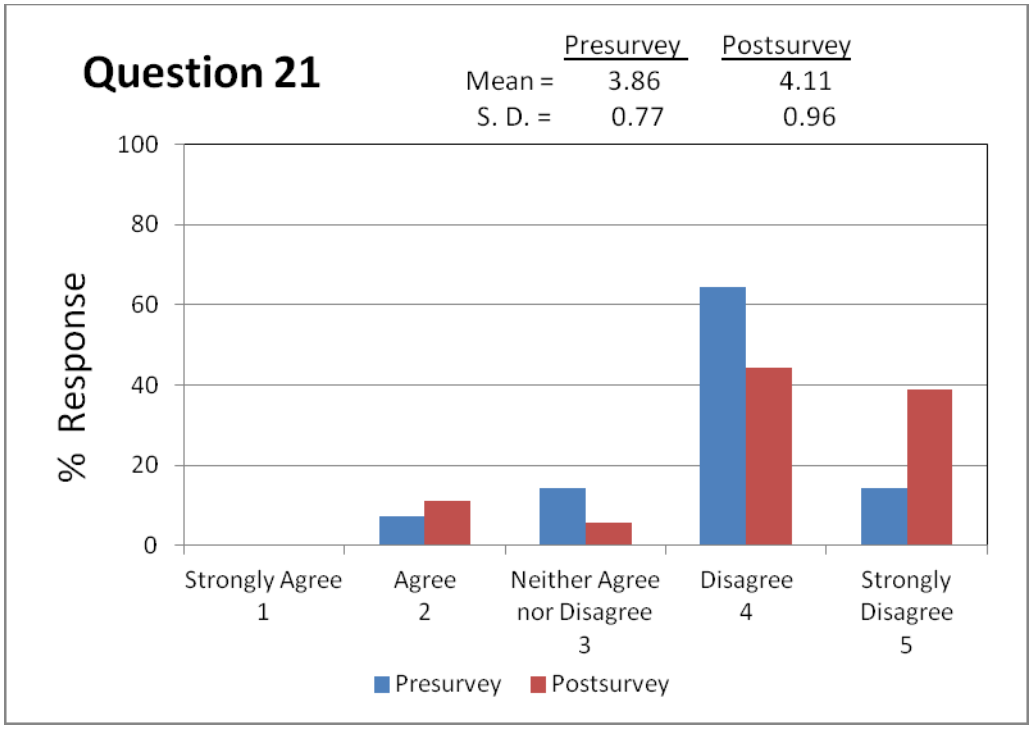


Question 21: It is okay to lower a student's grade based on disruptive behavior.

As illustrated in Figure 20, the comparison between the presurvey and postsurvey results indicated a mean shift from 3.86 to 4.11, but it was not statistically significant according to the Mann–Whitney U test. Although there was an increase in the percent of Strongly Disagrees in the postsurvey, the total of Disagrees and Strongly Disagrees (83.33%) was not very different from the total of Disagrees and Strongly Disagrees (78.57%) in the presurvey. There was a minimal perceptional shift in staff respondents concerning the acceptability of lowering a student grade based on disruptive behavior.

Figure 20

Okay to Lower Students' Grades Based on Disruptive Behavior

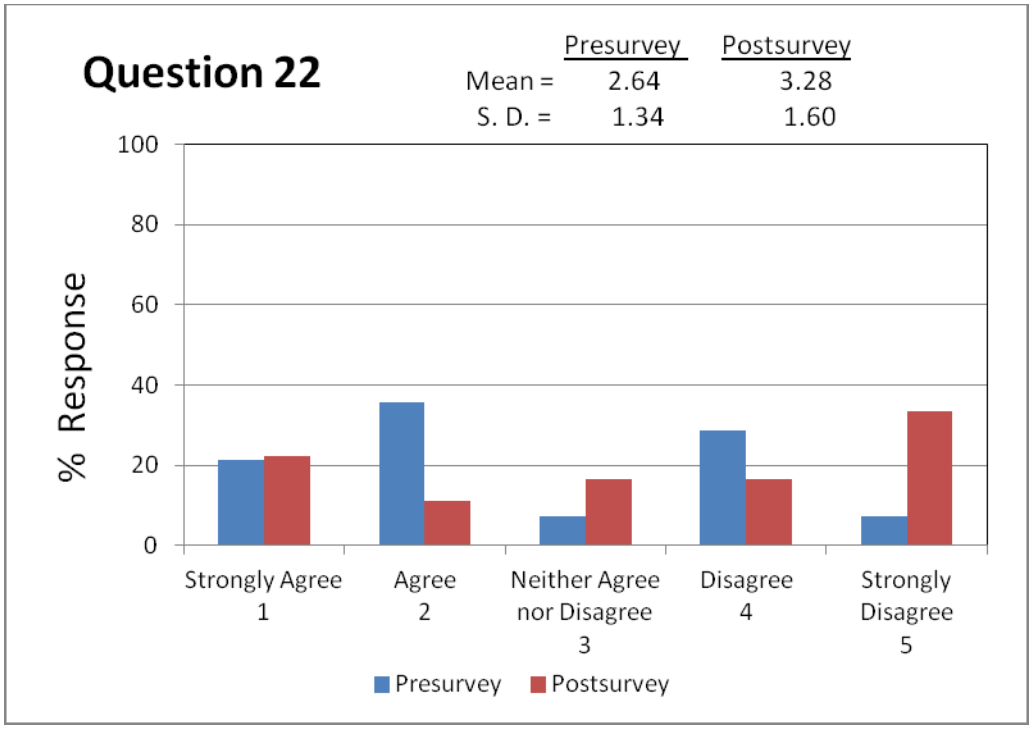


Question 22: If we did not grade student work, they would not do it.

As illustrated in Figure 21, the comparison between the presurvey and postsurvey results indicated a mean shift from 2.64 to 3.28, but it was not statistically significant according to the Mann–Whitney U test. There was an increase in the percent of Strongly Disagrees in the postsurvey. The total of Disagrees and Strongly Disagrees was 50%, as compared to 35.72% in the presurvey, yet it did not show statistical significance on the Likert scale. There was likely some perceptional shift in staff respondents concerning the contention if teachers did not grade student work, the students would not do it.

Figure 21

If We Did Not Grade Student Work, They Would Not Do It

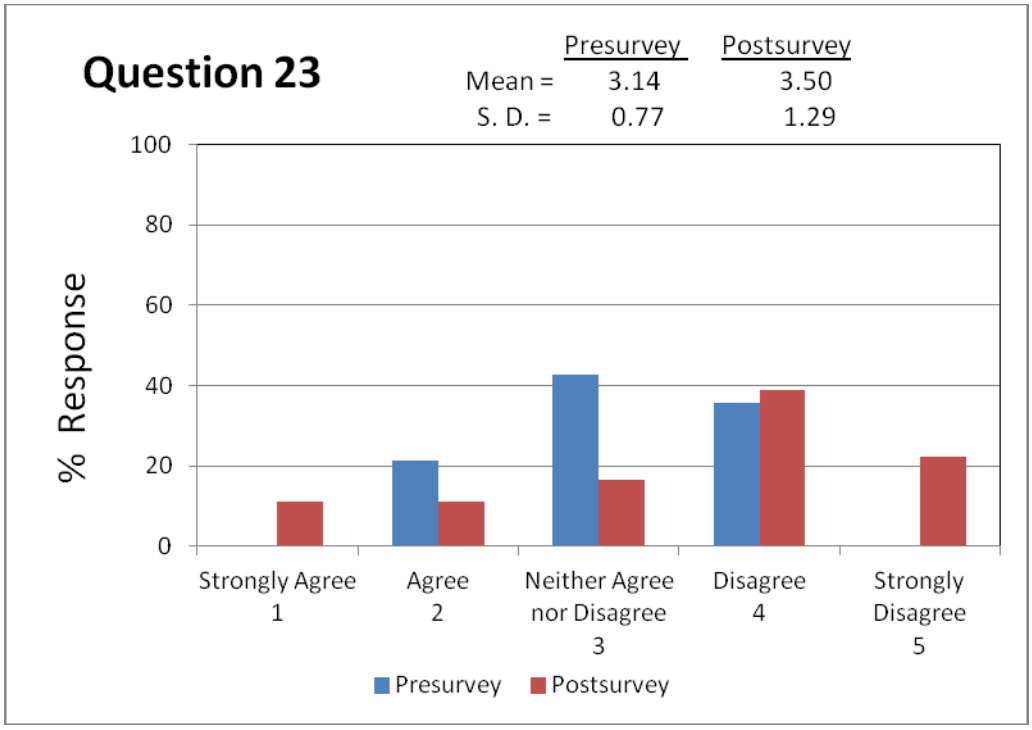


Question 23: Behavior and effort are important factors in academic achievement and should be reflected in students' grades.

As illustrated in Figure 22, the comparison between the presurvey and postsurvey results indicated a mean shift from 3.14 to 3.50, but it was not statistically significant according to the Mann–Whitney U test. There was an increase in the percent of Strongly Disagrees in the postsurvey. The total of Disagrees and Strongly Disagrees was 61.11%, as compared to 35.71% in the presurvey, yet it did not show statistical significance on the Likert scale. There was likely some perceptional shift in staff respondents concerning the contention behavior and effort are important factors in academic achievement and should be reflected in students' grades.

Figure 22

Behavior and Effort Should Be Reflected in Students' Grades



Chapter V

Introduction

This fifth and final chapter provides a summary and interpretation of the research findings of the study. Theoretical and practical implications, as well as recommendations for further research and conclusions, will also be delineated. The central research question posed for investigation and analysis in this study was “Can grading and assessment perceptions among middle level educators change through professional development?”

Significance of the Study

The significance of this study is multifaceted and points to the need for addressing issues, such as unsound and inequitable grading practices, resistance to change, low student motivation, the lack of quality assessment and feedback, and the need for professional development in middle level education. Current grading practices relying on a single letter or number value to communicate more than a student’s knowledge of a particular standard are no longer acceptable transmissions of academic achievement. Grades should accurately report subject comprehension so students, parents, teachers, administrators, future professors, and employers are able to depend upon the information for making decisions (Carey, 2001; Goodwin, 2011; O’Connor, 2007; Stiggins & Knight, 1998). When grading is too subjective, inconsistent among classrooms or schools within a district, or represents extracurricular factors, it is not a reliable source for reporting students’ comprehension of the subject matter (Balfanz, 2009; Guskey & Jung, 2012; Marzano, 2012; Stiggins, 2001). It is essential for educators to continually evaluate the practices and programs within the composition of a school in order to adequately meet

the needs of students. When administrators or teachers resist proposed change, it can cause unnecessary friction and delay vital adjustments (Dweck, 2006; DuFour, DuFour, Eaker, & Karhanek, 2010; Gullen, Gullen, & Erickson-Guy 2012; Guskey & Bailey, 2001). Student motivation is also negatively affected when poor grading practices are in place. This is dangerous because it causes students to disengage from their studies. Instead, they need to personally understand the value of investing in their own education through persistent effort (Chappuis, Commodore, & Stiggins, 2010; Dyrness & Dyrness, 2008). Quality assessments and feedback are key components to positively impacting student learning and must become an integral part of the educational process. When teachers frequently connect with students through assessment and feedback, they keep the students actively involved and prevent any surprises (Chappuis, 2009; Garrison, Chandler, & Ehringhaus, 2009; Stiggins, 2007; Vatterott, 2011).

Regular professional development is crucial for transforming ineffective grading and assessment practices and increasing student achievement. Perceptions in grading and assessment can change through strategic, research-based professional development opportunities. It is the primary venue for educators to learn how to apply the tools needed to make successful and lasting changes (Barr & Parrett, 2007; Brookhart, 2011; Dyb, 2012; Erickson, 2011; McDaniel, 2010). The importance of exploring best practices in grading and assessment, influencing change in the status quo, increasing student motivation, and offering recurrent, quality professional development pertaining to the education for middle level students cannot be overemphasized (Brookhart, 2011; Marzano, 2012; O'Connor, 2007; Stiggins & Knight, 1998).

Summary of Results

FRMS was selected as the location for this research study in which a presurvey (see Appendix A) and a postsurvey (see Appendix B) were conducted to assess the grading and assessment perceptions represented by the staff and to determine whether those perceptions changed following six professional development sessions. The staff members of FRMS who consented to participate answered 22 questions on each survey about grading and assessment practices. Dissenting and neutral responses in the presurvey for questions 2, 3, 4, 5, 7, 10, 11, 14, 15, 16, 17, 18, 19, and 20 were anticipated to shift to the Agree categories in the postsurvey. Assenting and neutral responses in the presurvey questions 6, 8, 9, 12, 13, 21, 22, and 23 were anticipated to shift to the Disagree categories in the postsurvey.

Of the 5-point Likert Scale questions in the study, four of them (2, 3, 15, and 16) revealed statistical significance by measuring patterns with enough difference in the comparison of both surveys to be considered statistically significant (any value $p = < .05$) according to the Mann–Whitney U test. In order to be determined statistically significant by the Mann–Whitney U test, the question must have shown a variance in response patterns. Since the Likert Scale is a continuum of perception, the different patterns may be interpreted as shifts in perceptions. The Mann–Whitney U test analyzes and figures ordinal and categorical data rather than numerical averages or scores (Tanner, 2012). It reports detectable patterns rather than shifts of averages. When research is conducted with small sample sizes and relatively few categories, a single change in response by one test subject can produce large shifts in averages but not necessarily impact the overall pattern. While cursory inspection of the averages may seem to indicate a significant

difference in responses, the Mann–Whitney U test shows statistical significance, according to the patterns of responses based on the calculation of their rankings.

Five questions (4, 5, 17, 19, and 21) showed little to no movement in perception because a majority of participants submitted the desired Likert responses in the presurvey. In question 4, 71.43% of the respondents were in the combined Agree and Strongly Agree categories in the presurvey and increased to 77.78% in the postsurvey. This showed a majority of the staff already understood the importance for grades to principally reflect achievement of standards. In question 5, 92.86% of the respondents were in the combined Agree or Strongly Agree categories in the presurvey and increased to 94.45% in the postsurvey. This showed nearly all of the staff already understood the importance for formative assessment to be frequent feedback on progress. In question 17, 78.57% of the respondents were in the combined Agree and Strongly Agree categories in the presurvey and decreased to 77.78% in the postsurvey. This showed a majority of the staff was already routinely involving students in formative assessments of their learning. In question 19, 78.57% of the respondents were in the combined Agree and Strongly Agree categories in the presurvey and decreased to 77.78% in the postsurvey. This showed a majority of the staff already believed high expectations for learning were communicated in the overall school setting. In question 21, 78.57% of the respondents were in the combined Disagree or Strongly Disagree categories in the presurvey and increased to 83.33% in the postsurvey. This showed a majority of the staff already understood the detriment of lowering students' grades based on disruptive behavior. This set of data is encouraging because it implies the FRMS staff was already aware and convinced of certain fundamental educational concepts. Issues of standards-based grading

(question 4), formative assessments (questions 5 and 17), environment for learning (question 19), and punitive grading (question 21) were understood and favorably marked as responses by the participants. The professional development provided by the in-service training was a tool reinforcing what most staff members' perceptions already exemplified.

The results of question 2 showed statistical significance, and Figure 1 (p. 63) illustrates the comparison between the patterns of responses in both surveys. The findings also demonstrated the staff was in favor of a school-wide grading and assessment policy. The combined Agree and Strongly Agree categories contained 71.43% of the respondents in the presurvey and increased to 94.44% in the postsurvey. This perception shift was the most substantial move toward implementing and maintaining sound grading and assessment practices within a school. Research indicates some districts leave the choice of grading and assessment practices entirely to the discretion of the individual teacher (Brookhart, 2011; Guskey & Bailey, 2001; Guskey, Swan, & Jung, 2011). When an overwhelming majority of a school's staff support the idea of common grading and assessment practices, the successful adoption and continuance of optimal measures is within reach.

The results of question 3 showed statistical significance, and Figure 2 (p. 64) illustrates the comparison between the patterns of responses in both surveys. The results also demonstrated the understanding of FRMS staff that giving zeros skews the mean (average) in calculating grades. The combined Agree and Strongly Agree categories contained 92.86% of the respondents in the presurvey and increased to 100% in the postsurvey. This showed, after professional development sessions addressing the negative

impacts of recording zeros, the entire group was convinced of the logic behind the mathematical inaccuracy. A grade is skewed and the true measurement of a student's knowledge is unattainable when zeros are used (Wormeli, 2006). Practices such as recording zeros are widely employed but do not sufficiently reflect a student's progress and, furthermore, may present inaccurate evidence falsely representing learning deficiencies (Brookhart, 2011; Guskey & Bailey, 2001). A zero score is an invalid reporting method that does not accurately describe a student's comprehension of subject material and should not be used for transmitting academic placement (Marzano, 2006; O'Connor & Wormeli, 2011). The results of the postsurvey on question 3 are encouraging because when teachers begin to understand how using zeros profoundly affects a student's grade, they are only a few steps from making significant changes in their grading practice. Belief precedes action; if teachers believe using zeros has a negative impact, they are more likely to choose alternate ways to communicate with parents about missing assignments.

The results of question 15 showed statistical significance, and Figure 14 (p. 76) illustrates the comparison between the patterns of responses in both surveys. The findings also demonstrated that the staff understood the purpose of summative assessment to be a periodic report of students' achievement on standards. The Agree category contained 57.15% of the respondents in the presurvey and increased to 88.88% in the combined Agree and Strongly Agree categories in the postsurvey. This has at least two implications. First, teachers may have come to a better understanding of the purpose of tests and other summative assessments as reporting measures of a student's understanding of predetermined educational standards rather than an arbitrary end-of-unit

activity. Second, teachers may have learned how to distinguish the features and purposes between formative and summative assessments. Teachers should give valuable and applicable feedback to the students throughout the course of instruction, which are assessments for learning (formative assessments) before giving high-stakes tests (summative assessments) or assessments of learning (Chappuis, 2009; Fisher & Frey, 2007). Either way, the magnitude of the shift means teachers benefitted from instruction during the professional development sessions pertaining to summative assessments.

The results of question 16 showed statistical significance, and Figure 15 (p. 77) illustrates the comparison between the patterns of responses in both surveys. The results also demonstrated the staff understood recent evidence is more valuable than an average of points over time when figuring summative grades. The Agree category contained 42.86% of the respondents in the presurvey and increased to 66.67% in the combined Agree and Strongly Agree categories in the postsurvey. This also showed the professional development sessions provided beneficial instruction increasing the teachers' knowledge base and understanding of the purpose of summative assessments. Teachers were challenged to consider data claiming students should not be punished or given a lower grade because they grasped a concept during week 3 instead of week 1. Teachers may apply various strategies while constantly keeping in mind the ultimate goal is for students to meet the learning expectations of a particular standard. Again, the positive shift in this question signified teachers gained perspective and guidance during the in-service training on summative assessments.

Three questions (5, 14, and 19) displayed pattern variance close to statistical significance; however, they did not show enough pattern variance to measure statistically

significant according to the Mann–Whitney U test. In question 5 (assent desired), 92.86% of the respondents were in the combined Agree and Strongly Agree categories in the presurvey and increased to 94.45% in the postsurvey. In question 14, 28.57% of the respondents were in the Agree category in the presurvey and increased to 66.66% in the combined Agree and Strongly Agree categories in the postsurvey. In question 19, 78.57% of the respondents were in the combined Agree and Strongly Agree categories in the presurvey and decreased to 77.78% in the postsurvey. While comparison of the results in questions 5 and 19 showed little percentage fluctuations, they were still near statistical significance according to the Mann–Whitney U test. This is the result of the ordinal nature of the Mann–Whitney U test. A percentage shift calculation regards the Likert scores of 1 and 2 as similar, positive perceptions as described by Strongly Agree or Agree, while the Mann–Whitney U test treats them as independent or unrelated categories (Tanner, 2012). Question 14 also showed a lot of difference in the pattern, consequently near the statistically significant determination. A high-percentage shift fails to acknowledge the large number of neutral scores. When those neutrals are taken into account by the Mann–Whitney U test, the presurvey and postsurvey difference is not as extreme as the percentage shift implies.

The themes addressed in the survey questions and extracted for the research study included grading policies, grading scales and reports, grades as a reflection of achievement on standards, grade pollution, ability grouping, summative assessments, formative assessments, grading information systems, high expectations for learning, and the inclusion of behavior and effort within a grade. Some showed minimal shifts in responses from the presurvey to the postsurvey, while others showed substantial shifts.

There are numerous possible explanations to be delineated for these results, expounding on the data and the perceptions of the participating respondents.

Grading Policies

Questions 2 and 9 explored the theme of a school-wide policy versus individual teachers' grading practices. Whereas the results in question 2 (assent desired) showed 94.44% of the respondents in the postsurvey would support the idea of a school-wide grading and assessment policy, only 50% of the respondents in question 9 (dissent desired) said teachers should not be able to adopt whatever grading practices work for them (see Figure 1, p. 63 and Figure 8, p. 70). Question 9 had 21.43% of the respondents in the presurvey who disagreed with teachers individualizing grading practices, and although there was an increase to 50% in the postsurvey, it revealed an incongruity with responses to question 2 addressing uniform grading practices through the establishment of a school-wide policy. FRMS staff showed a willingness to embrace a school-wide policy; however, there was a detectable inconsistency when contrasted to the responses of question 9. Perhaps this is because there were diverse interpretations of the term grading practices in question 9. For some, it may have meant there would be only one way to report a grade. Perhaps they were hesitant to give up autonomy in determining how to assess and report student achievement because unknown variables, such as homework, ratios of formative to summative assessments, and extra credit allowance, were not addressed in the same question. The data clearly proved FRMS is ready to move in the direction of a school-wide grading and assessment policy; however, it is unclear how easy the process might be. The independent minds of professionals certainly contribute to the equation and may require compromises until the full scope of the impact of grading

practices on student achievement is realized. Grading and assessment practices have great influence, positive and negative, on the short-term and long-term educational success of students. They have the power to affect student motivation, effort, and achievement (Bonesronning, 2004; Guskey, 2011). Individual schools and districts that have not agreed on a uniform grading philosophy perpetuate inconsistency throughout the program (Guskey & Jung, 2012; Marzano, 2010). FRMS seems to have appreciable potential for aligning grading and assessment practices with its staff and eventually approving a uniform, school-wide grading policy. The friction may surface when the policy on the table is not agreeable to some of the teachers. It is the hope for the FRMS staff to strive for the goal of a uniform grading policy through communication, collaboration, and compromise.

Grading Scales and Reports

Questions 3, 6, and 11 exposed the common theme of what is considered fair and accurate when reporting grades. In question 3 (assent desired), 92.86% of the respondents were in the combined Agree and Strongly Agree categories in the presurvey and increased to 100% in the postsurvey. This showed all of the teachers in the group understood how giving zeros skews the average score when calculating grades and the increase was most likely due to the professional development sessions. In question 6 (dissent desired), 21.43% of the respondents were in the Disagree category in the presurvey and increased to 33.33% in the combined Disagree and Strongly Disagree categories in the postsurvey. This showed a minority of teachers in the group understood the problems with basing grades on a mean of points accumulated in a traditional scale. In question 11 (assent desired), 57.15% of the respondents were in the Agree category

and increased to 72.22% in the combined Agree and Strongly Agree categories in the postsurvey. This showed a majority of the group understood using a grade of incomplete is preferable to penalizing late work or giving zeros or Fs, and this increase was most likely due to the in-service training. The widely accepted practice of assigning zeros or Fs was challenged in these questions. The hopeful expectation was for teachers to understand the considerable impact such practices have on the final grade and ultimately the overall educational experience of a student. Research indicates an A, which is 90%–100%, has a much smaller range in scale compared to an F, which is 0%–59%, and assigning an F grade reaps the same ill effects as assigning zeros and should be considered an inappropriate measure for grading (Gerke, 2007; Schimmer, 2012). The practice of giving zeros should be abandoned (Brookhart, 2009; Wormeli, 2006), and in its place could be alternatives, such as rubric scales (Wormeli, 2006), a bottom line of 60 (Schimmer, 2012), assignment and test retakes (Dueck, 2011; Gordon & Fay, 2010; Wormeli, 2011), and notations of incomplete (Reeves, 2004). Until teachers appreciate the significance of grading practices on student understanding, motivation, and achievement, they will continue to use practices based on punishment and revenge (Erickson, 2001; O'Connor, 2007). Grades are intended to assess and promote learning, and there are many ways to calculate and report them fairly and accurately without giving zeros (McMunn, Schenck, & McColskey, 2003; Popham, 2008; Winger, 2005).

Several professional journal readings addressing the harmful impacts of punitive grading were incorporated into the in-service training. If teachers were genuinely interested in the accurate report of student achievement, they would not continue the practice of assigning zeros for any reason. The inconsistency lies with the fact that

although all the FRMS teachers in the research group understood the mathematical inaccuracy of using a zero (question 3), only one third were willing to scrutinize the validity of the traditional grading scale (question 6), and more than one fourth were still not convinced using a grade of incomplete is preferable to penalizing late work by giving zeros or Fs. Perhaps this indicates unwillingness on the part of some teachers to do the extra work it might take to reenter grades into the grade book when incomplete assignments are finally submitted, even while knowing the research. Maybe it shows the tendency for some teachers to hold ardently to traditional practices simply because they are comfortable with the routine. Some educators believe current practices accurately reflect how the real world works and operates (Balfanz, 2009; Erickson, 2001), and they are hesitant to discontinue a holistic grading approach because of the ardent belief students must be held accountable for their choices (Balfanz, 2009). Educators may also express resistance to change because a large number of current practices have deep roots in educational traditions. Teachers are more comfortable with past methods than contemporary or progressive approaches (Guskey & Bailey, 2001; Guskey, 2006). It may be deduced that the mere understanding of a concept or principle does not necessarily transfer to a change in practice or tradition for some. Therein lies the quandary of teachers' perceptions failing to transfer to classroom practices. Nonetheless, positive shifts in these questions illustrated the constructive instruction during the professional development sessions provided during the in-service training.

Grades as a Reflection of Achievement of Standards

Questions 4 and 7 presented the topic of grades serving as the principal reflection of a student's achievement on predetermined standards or class objectives. The desired

result was for a majority of teachers to agree with the suppositions made in the surveys. In question 4 (assent desired), 71.43% of the respondents were in the combined Agree or Strongly Agree categories in the presurvey and increased to 77.78% in the postsurvey. This meant most of the teachers in the group believed grades should principally reflect achievement of standards; however, the professional development did not significantly increase the majority of respondents (Neutral and Disagree movement) or move enough of them into the Strongly Agree category from the Agree category. In question 7 (assent desired), 28.57% of the respondents were in the combined Agree and Strongly Agree categories in the presurvey and increased to 44.44% in the postsurvey. This indicated less than half of the teachers believed it was okay to give students an A grade when they have mastered a course objective but not completed all homework or class work.

When academic achievement is the only factor indicated in the letter grade, it provides a true depiction of the students' understanding of the course standard (Carey, 2001; Stiggins & Knight, 1998). The intervention did not generate a large perception change in these questions (see Figure 3, page 65 and Figure 6, page 68). The inconsistency lies with the fact that although three fourths of the staff believed grades should reflect achievement of standards, less than half were willing to critically examine current practices of including homework and class work within a grade. This may indicate reluctance on the part of some teachers to change traditional practices of including homework within the grade, even when they realized the research and logic surrounding the subject supported otherwise. Again, the conjecture is the mere understanding of a concept or principle does not necessarily transfer to a change in

practice or tradition for some. Once more, the dilemma is posed regarding how to move teacher perception to teacher practice.

Grade Pollution

Questions 8 and 10 addressed the common theme of grade pollution and challenged the teachers at FRMS to consider the potentially damaging effects of including more than students' knowledge in a grade. In question 8 (dissent desired), 35.71% of the respondents were in the combined Disagree and Strongly Disagree categories in the presurvey and increased to 44.44% in the postsurvey. This showed more than half of the teachers were still not convinced after the professional development sessions that penalizing students for late work is an unacceptable practice. In question 10 (assent desired), 28.58% of the respondents were in the combined Agree and Strongly Agree categories in the presurvey and increased to 55.56% in the postsurvey. While the percentage shift was substantial in the assent categories, the Likert scores ignore the large numbers in the Neither Agree nor Disagree and Disagree categories and the relative similarities in the presurvey and postsurvey categories. The Mann–Whitney U test analyzes all categories, and the overall amount of change was not statistically significant (see Figure 9, page 71).

The postsurvey results indicated about half of the teachers in the group understood the problems with continuing the faulty practice of allowing extra credit. Anything not related to the achievement level, such as late work or extra credit, should not be recorded in the achievement section of the report card (O'Connor, 2007; Roorda, 2008). Punishing students for failure to turn in homework or turning in late homework represents an ineffective grading practice (Erickson, 2001; Wormeli, 2006) and distorts

the point of reference that should be a student's level of understanding of a particular objective or standard (O'Connor, 2007). Results for question 8 demonstrated an almost imperceptible shift after the professional development sessions, which meant the majority of teachers still believed in penalizing students for missing deadlines. The in-service training did little to change the minds of the participants. Again, perhaps this is because teachers are resistant to changing traditional practices. As mentioned in one of the session videos, teachers fundamentally want to exact revenge on students for not adhering to the rules and expectations of the class. Maybe teachers in the group were not receptive to the allegations made during the intervention. The opinions of teachers play too big of a role when delivering course grades (Guskey & Bailey, 2001), and the conventional practice of recording typical letter grades for student work reflects a reliance on professional judgment based on little objectivity and too much subjectivity (Aronson, 2008; Guskey & Bailey, 2001; Guskey, 2006). Question 10 showed a much more substantial shift in perception pertaining to the use of extra credit. Whereas, only half of the group in the postsurvey believed extra credit should not be given, there was still an observable adjustment made, illustrating the positive impact of the professional development on a portion of the participants. Nevertheless, this showed the difficulty of transforming teachers' perceptions and, more importantly, transferring change into the classroom through grading and assessment practices.

Ability Grouping

Question 13 tackled the concept of ability grouping. In question 13 (dissent desired), 39.86% of the respondents were in the combined Disagree and Strongly Disagree categories in the presurvey and increased to 55.56% in the postsurvey. This

showed more than half of the staff understood grades should not be used to group students according to ability. Ability grouping is a practice teachers have embraced for a long time. For many years, teachers have considered grades as an indicator of a student's ability, whereas grades are not intended to place kids in high, medium, and low categories. Grades should not be a tactic for comparison and contrast. Grades have the power to influence how students feel about themselves and their future ability to learn (Dyrness & Dyrness, 2008; Schimmer, 2012), and when students are given a failing label, they become susceptible to an attitude of defeat and consequently pull back from the educational arena; rather than increasing desire and effort, the failure impedes them (McMillen, 2012). There was a positive shift in the results of the survey responses, which pointed to the positive impact of the professional development on some of the participants.

Summative Assessments

Questions 14, 15, and 16 presented the common theme of summative assessments. In question 14 (assent desired), 28.57% of the respondents were in the Agree category in the presurvey and increased to 66.66% in the combined Agree and Strongly Agree categories in the postsurvey. Even though the percentage shift is extreme, the Mann–Whitney U test showed a near significant difference. It scored a .062, which is close but does not meet the criteria of .050 for significance. One contributing factor could be the small sample size for the surveys. When there is a statistical test on a small sample size, the pattern change required to show statistical significance is greater in order to prevent a reading of random fluctuation as a significant difference. In question 15 (assent desired), 57.15% of the respondents were in the Agree category in the presurvey and increased to

88.88% in the combined Agree and Strongly Agree categories in the postsurvey. In question 16 (assent desired), 42.86% of the respondents were in the combined Agree and Strongly Agree categories in the presurvey and increased to 66.67% in the postsurvey. Teachers should give valuable and applicable feedback to the students throughout the course of instruction, which are assessments for learning (formative assessments) before giving high-stakes tests (summative assessments), or assessments of learning (Chappuis, 2009; Fisher & Frey, 2007).

Summative assessments provide an opportunity for students to demonstrate their knowledge (Chappuis, Stiggings, Chappuis, & Arter, 2012) and all they have learned in preparation (Davies, 2009; Fleenor, Lamb, Anton, & Donen, 2011; Rom, 2011). A quality summative assessment should address the concepts previously taught in class and be weighed according to the degree of difficulty and time spent on the material (Salend, 2011). By gaining an understanding of the purpose of quality summative assessment, teachers are on the right track to improving the educational experience for students. There were substantial shifts in all of these questions. The positive movement points to the impact of the professional development sessions on teachers' understanding of the purpose of summative assessment to report students' achievement of standards.

Formative Assessments

Questions 5 and 17 identified the issue of formative assessment. In question 5 (assent desired), 92.86% of the respondents were in the combined Agree and Strongly Agree categories in the presurvey and increased to 94.45% in the postsurvey. In question 17 (assent desired), 78.57% of the respondents were in the combined Agree and Strongly Agree categories in the postsurvey and decreased to 77.78% in the postsurvey. Formative

assessments increase academic achievement and, unlike summative assessments, are given with plenty of time for students and teachers to make changes to improve knowledge and understanding (Varlas, 2012). Formative assessment is an invaluable tool offering the most benefits for teachers and students when administered frequently (Brookhart, 2009; Chappuis, Stiggins, Chappuis, & Arter, 2012; Popham, 2008), and it should assist the teacher in determining instructional adjustments with the purpose of improving student achievement (Brookhart, 2009; Davies, 2009; Fisher & Frey, 2007; Willhoft, 2012). Formative assessments provide the stage for students to learn standards and course objectives through a regular instruction and feedback process (Doubet, 2012). Formative assessments give teachers immediate results to discern the most current issues with student understanding (Bartlett, 2012). The goals of formative assessment are to improve scores, understanding, confidence, and quality learning for students (McMunn, Schenck, & McColskey, 2003; Doubet, 2012; Stiggins, 2007), and huge strides are gained when students are properly measured (Stiggins, 2006). The data showed FRMS teachers already understood the purposes and features of formative assessments and were incorporating them into the classroom.

Grading Information Systems

Questions 18 and 20 addressed the issue of the use of grading information systems as a primary tool for reporting grades. In question 18 (assent desired), 7.69% of the respondents were in the Agree category in the presurvey and increased to 22.22% in the postsurvey. The percentage shift here was due to the selection of only two Likert scales within the comparison. Again, the Mann–Whitney U test looks at all categories and the amount of difference shown from the presurvey to the postsurvey in the patterns

of responses according to the calculation of their rankings. In question 20 (assent desired), 71.43% of the respondents were in the combined Agree and Strongly Agree categories and decreased to 61.11% in the postsurvey. Question 18 referred to the difficulty of translating various scores and notations into a district-wide grading software program (e.g., PowerSchool). Most grading software programs do not include features for reporting student scores with best practices in grading and assessment in mind. The small percentage of participants who agreed with this supposition indicated the teachers did not understand the subject of grading systems or the professional development did little to explain and clarify or both. Responses to question 20 can vary based on a teacher's subject. Math teachers would readily agree student scores must be recorded weekly, if not daily. Language Arts and Social Studies teachers would claim grades need only to be recorded biweekly.

Research indicated another source for learning new grading techniques is computer software. Some software programs allow a teacher to mark only the grades, yet others give options allowing teachers flexibility in recording procedures, such as breaking down a unit into quizzes (formative assessment), homework, and midterms (Gullen, Gullen, & Erickson-Guy, 2012; Marzano, 2006). Some programs are easier to use than others and may provide teachers the ability to manipulate their grading procedures within the software. An effective grading software system will go beyond the mere spreadsheet calculation of averages, allowing teachers to record both formative and summative assessments as they relate to standards (Marzano, 2006).

High Expectations for Learning

Question 19 posed the inquiry regarding the school's expectations. In question 19 (assent desired), 78.57% of the respondents were in the combined Agree and Strongly Agree categories in the presurvey and decreased to 77.78% in the postsurvey. FRMS was decidedly successful in communicating high expectations for learning in the overall school setting. According to the presurvey, teachers already clearly believed their administration and staff did a good job of establishing a culture promoting and rewarding student learning. This is why there was very little shift in perception from the presurvey to the postsurvey as shown in the data. In the Strongly Agree category, there was a shift from 7.14% to 50%, which is especially noteworthy. This evidence is favorable to the possibility of the future implementation of a school-wide grading and assessment policy because there is a foundation of communication. Research indicated success in the implementation of uniform grading practices is achievable when there is staff support (Gullen, Gullen, & Erickson-Guy, 2012). A supportive staff capable of open and constructive communication is a great first step toward initiating a unified approach and creating guidelines for common practices.

Inclusion of Behavior and Effort Within the Grade

Questions 12, 21, 22, and 23 addressed the issue of including behavior and effort within a student's grade. In question 12 (dissent desired), 42.86% of the respondents were in the Disagree category in the presurvey and increased to 44.44% in the combined Disagree and Strongly Disagree categories in the postsurvey. This demonstrated there was no perceptional shift and showed the majority of the participants were not convinced a student's behavior and work ethic should be separated from the achievement grade.

Teachers are still uncomfortable not including behavior and work ethic within the grade. This could be because the grading systems do not allow for it. It may be the inherent flaws in the design of the report card. Teachers say it is difficult to convince parents to care about the behavior and engagement reports as much as they care about the academic reports. Some teachers are concerned about the ability to motivate students without using grades as a hammer. In question 21 (dissent desired), 78.57% of the respondents were in the combined Disagree and Strongly Disagree categories in the presurvey and increased to 83.33% in the postsurvey. This demonstrated a minimal shift and showed the majority of the group did not believe it was acceptable to lower a student's grade based on disruptive behavior. In question 22 (dissent desired), 35.72% of the respondents were in the combined Disagree and Strongly Disagree categories in the presurvey and increased to 50% in the postsurvey. This demonstrated more movement than the previous two questions and showed half of the teachers believed if they did not grade students' work, the students would not do it. The goal of the professional development sessions was to stress the importance of emphasizing concept learning and promoting the intrinsic value of education as opposed to merely striving for a good grade. In question 23 (dissent desired), 35.71% of the respondents were in the Disagree category in the presurvey and increased to 61.11% in the combined Disagree and Strongly Disagree categories in the postsurvey. This demonstrated the most substantial shift of all four questions and showed the majority of the group believed, after the professional development sessions, that behavior and effort should not be reflected in a student's grade.

It is crucial to differentiate the components of a grade, instead of lumping them all together, so academic information is uncomplicated and clearly reported (Goodwin,

2011; O'Connor, 2007). The complexity of a grade on a typical report card makes understanding difficult and accurate assessment unattainable. A grade may incorporate aptitude, achievement, effort, compliance, and attitude within a single number or letter (McMillan, Myran, & Workman, 2002; McDaniel, 2010; Shanahan, 2011; Stiggins, 2001). As long as schools continue this practice, it poses a dilemma because the grade is not pure but rather reflects numerous variables impossible to differentiate and clearly understand (Carey, 2001; Stiggins & Knight, 1998). While behavior issues are important and need to be communicated with parents and students, recorded grades should only reflect a student's knowledge of concepts and academic progress (Balfanz, 2009; Grimes, 2010; Marzano & Heflebower, 2011; Rom, 2011). It is vital for academic grades to be separated from effort and behavior, and teachers and administration can still require a high level of responsibility on the part of students (Arter & Chappuis, 2006; O'Connor, 2007; Haselhuhn, Al-Mabuk, Gabriele, Groen, & Galloway, 2007).

There were positive shifts in all four questions, which points to the constructive impact of the professional development sessions. The inconsistency lies with the fact that although most of the teachers understood it was not an acceptable practice to lower a student's grade due to disruptive behavior (question 21), less than half were willing to scrutinize the practice of including behavior and work ethic within the achievement section of the grade (question 12). Yet, some were persuaded during the in-service training to separate behavior and effort from academic achievement (question 23). This may simply indicate a misunderstanding of a survey question. It may also imply reluctance on the part of some teachers to discontinue grading practices they deem necessary to punish or motivate students. The professional development sessions strove

to highlight the importance of separating behavior, work ethic, effort, and all other nonacademic issues from the achievement portion of a student's grade. Again, the conclusion is the mere understanding of a concept or principle does not necessitate the transfer to a change in practice or tradition for some. The dilemma is posed regarding how to move teacher perception to teacher practice.

Impact of Limitation

A limiting factor of this study was that the research test group consisted of one rural school with a relatively small teaching staff of 22 to 25 teachers. Another limitation was the three-month window between the presurvey and postsurvey. This may have not allowed sufficient time for teachers to explore grading and assessment practices through a trial and error process. The fact that only 14 teachers took the presurvey and 18 teachers took the postsurvey limited consistent outcomes of the research and translated into reporting four respondents in the postsurvey who were not included in the first set of data. Prior to this research study, it was perceived FRMS had limited exposure to grading and assessment training; however, after further investigation, it was realized the teachers had previous instruction in grading and assessment practices. FRMS has received several national recognitions, and the staff had already begun independently investigating and incorporating sound grading and assessment practices within the school, although their practice was on a limited basis and within certain pockets of the staff.

Conclusion

This research study embarked on the task of selecting a workable target group (FRMS), measuring the consenting participants' perceptions on grading and assessment through a presurvey, conducting relevant intervention through professional development

sessions, and finally determining the participants' perceptions on the same questions through a postsurvey. The evaluation of the research study conducted with the FRMS staff can be delineated by reviewing the results of the presurvey and postsurvey and comparing the differences in responses to draw reasonable deductions. Based on the positive shifts in percentages for 19 of the 22 survey questions and the determination of statistical significance on four of the questions, FRMS is on its way to unifying grading and assessment practices. By offering more frequently scheduled, quality professional development, they could achieve the goal of a school-wide policy. Continual opportunities for professional development and collaboration are vital to the introduction and maintenance of effective grading and assessment practices (Dyb, 2012; Roorda, 2008). If the FRMS teachers had ample time to explore best practices, they could determine which methods worked for them and would be the best choices to implement throughout the school. Additional assessment training should be made available for the staff at FRMS over the next several years in order for there to be a complete fundamental shift in perception and practice regarding grading and assessment.

The most important themes pertaining to grading and assessment were addressed in the survey questions. Perceptions of the concept of a school-wide grading policy versus individual grading practices were measured in questions 2 and 9. Perceptions of assigning zeros, Fs, and using traditional grading scales were measured in questions 3, 6, and 11. Perceptions of grades reflecting students' achievement of standards as opposed to homework or class work were measured in questions 4 and 7. Perceptions of penalizing students' late work and offering extra credit were measured in questions 8 and 10. Perceptions of sorting students based on ability were measured in question 13.

Perceptions of summative assessment were measured in questions 14, 15, and 16. Perceptions of formative assessments were measured in questions 5 and 17. Perceptions of the impact of grading information systems were measured in questions 18 and 20. Perceptions of expectations for learning in the school setting were measured in question 19. Perceptions of the issue of including effort, behavior, and work ethic within a grade were measured in questions 12, 21, 22, and 23. As the data and figures displayed, there were four questions with statistical significance, three questions close to statistical significance, and five questions that showed a majority of the participants were at the desired Likert response in the presurvey. Most of the postsurvey responses demonstrated percentage increases in responses, indicating positive perception shifts. These measurements occurred with only six professional development sessions and a three-month time span between surveys.

The first step in the introduction and eventual implementation of new concepts or practices is to evaluate and understand the current ideologies represented by the target group. Once this is determined, the next step is to strategize the best methods for challenging the dogma of the members of said group. The agreed-upon intervention would be conducted followed by a second assessment of the participants' perceptions to ascertain whether any desired shifts were made. If the results are positive, plans to discuss the application of new ideas would be scheduled and the ultimate goal of a perception change translating to a practice change would be achieved. This was the essence of the research study, and it is the ever-present goal of middle level education. Educating middle school students presents issues that must be continually assessed according to value and productivity. If standards for learning are the proven tools for

measurement, then practices unsupportive of those standards must be subject to appraisal and possible rejection. This study offered a small-scale version of what middle schools across the country could achieve if professional development became their mantra. New methods and strategies for grading and assessment are introduced, practiced, evaluated and modified through professional development. Purposeful community is developed through professional development. The entire educational process is enhanced through regular, quality professional development. The final conclusion of the results of this study points to the constructive potential for the impact professional development has on the perceptions and eventual grading and assessment practices of middle level educators.

Recommendations for Further Research

This research study engaged in the issue of the influence of professional development on the perceptions of middle school educators regarding grading and assessment. Grades exemplify teachers' central philosophies about education and its purposes (Gullen, Gullen, & Erickson-Guy, 2012), and there is a wide spectrum of beliefs on grading practices (Scherer, 2011). The survey and subsequent professional development sessions endeavored to challenge some generally accepted beliefs about grading and assessment. It disputed the value of traditional practices and presented viable alternatives based on research and results. After conducting the study, many ideas surfaced about the continuance of similar research. Recommendations for additional research include adding a qualitative study, selecting varied and larger test groups, allowing more time for professional development and practice, assessing specific and unique needs of the test group, and using an additional testing instrument.

The first recommendation relates to the scope of the study. This study examined perceptions according to two variables, the presurvey and the postsurvey. The inclusion of practices should be considered and would require questionnaires. A mixed-method research study would add substance through the use of a qualitative research design to investigate teachers' individual grading and assessment practices.

A second recommendation relates to the target population. This study selected a relatively small staff in one rural middle school. It is a possibility the findings may be an anomaly specific to this setting. Future studies may involve more than one school, as well as schools with larger staffs in varied locations.

A third recommendation relates to the amount of time given for the professional development intervention. Six sessions were presented to FRMS through an in-service training module. Further research may schedule biweekly professional development opportunities (book studies, case studies, webinars, etc.) to provide an exhaustive set of resources pertaining to grading and assessment. Online instruction could also be conducted asynchronously or synchronously to supply more time for lessons and practice.

A fourth recommendation relates to the duration of the practice time. This study allowed the staff three months to apply best practices in grading and assessment learned in the professional development sessions. For future research, a longitudinal study should be conducted allocating at least one year for practice and implementation, and changes could be tracked after the third year to determine the sustainability and maintenance of those changes.

A fifth recommendation relates to the investigation of the specific and unique needs of the target group. This study used a predetermined set of survey questions. A

future study should conduct a needs assessment prior to any planned intervention to properly determine the distinctive desires and necessities with regard to grading and assessment within the test group. This would prevent unnecessary repetition or duplication of information they already know.

A sixth recommendation relates to the use of an additional testing instrument. This study applied the Mann–Whitney U test to determine statistical significance based on the comparison of response patterns. For this study, the ideal method would have been a paired *t*-test, however anonymity could not be guaranteed as required. There should be an investigation into how to issue random identification cards or numbers to ensure anonymity but still pair a single individual with his or her responses.

Implications for Professional Practice

The data showed professional development had a positive impact on the FRMS staff's perceptions about grading and assessment practices. In nearly every question, there was movement toward the desired Likert score. Even though only four questions were labeled statistically significant and three were close to statistical significance, the data revealed a heightened revelation of the teachers' growth in their perceptions. The goal of this research was to positively influence the perceptions of the participants through professional development training sessions to understand the detrimental effects of past and current grading and assessment practices and the benefits of alternative strategies. Of course, the ultimate goal would have been to convince 100% of the participants to answer according to the desired Likert scores in the postsurvey, which was achieved on only one, question 3. This showed it is possible to impact every participant in some way, and professional development is the primary component for making these

accomplishments. The theoretical implications for professional practice point to the need for strategic, consistent, and continuous professional development. Staff training is necessary to impart a vision; it sets the stage with concepts, strategies, and tools; it creates a plan for putting the learned skills into practice and enhancing the purposeful community. The practical implications point to the success of professional development. This study demonstrated the success of professional development for a staff, even on a small scale. When schools or districts decide to initiate and implement change, training through a planned professional development module is the most ideal strategy for presenting the proposal (Dyb, 2012; Gullen, Gullen, & Erickson-Guy, 2012).

Application

Professional development offers the best chance for the successful implementation, maintenance, and sustainability of changes in the fundamental approaches to grading and assessment practices in education, forever impacting the lives of our students (Dyb, 2012; Gullen, Gullen, & Erickson-Guy, 2012; McDaniel, 2010). In order to implement practical applications of grading and assessment best practices in any school setting, the administration and teachers must commit to a strategic professional development plan. Continual opportunities for professional development and collaboration are vital to the introduction and maintenance of effective grading and assessment practices (Dyb, 2012; Roorda, 2008). To achieve a purposeful community through the implementation of grading and assessment, the staff will require professional development opportunities with extended time for application. Districts and schools have been able to achieve the desired professional development agenda through early release times, prep-hour meetings, and regularly scheduled department meetings. Funding these

initiatives on a school-wide or district-wide basis must be a priority when constructing the annual budget. The district or school should commit to funding these efforts through Title One money or general money in order to convert improved teachers' perceptions of grading and assessment to applied practices in the classroom.

Recruiting key people to assist staff when implementing or applying a new policy is necessary. Patience is a virtue and when the stakeholders understand the application process may take three to six years to be fully assimilated into their schedules and daily practices, they will be less likely to criticize and more likely to modify along the way. Administrators would do well to communicate clearly to their staff throughout the application process, assuring the freedom to experiment and adjust without fear of critical evaluation. The creative and practical options for new-and-improved grading practices are countless, providing schools with tools for achieving a greater degree of student success within the educational process (Erickson, 2001; O'Conner, 2007). Continuous professional learning leads to better classroom practices and success (Teague & Anfara, 2012) and guarantees grades become a true representation of student knowledge and achievement (Aronson, 2008; Moss & Brookhart, 2009).

The applications of this research, as implied by the survey questions, include the presentation and assimilation of best practices in grading and assessment through professional development. Some educational approaches performed every day in classrooms throughout the country are not working (O'Connor, 2007). To learn and intentionally apply new strategies takes courage and commitment. The research highlights the most prominent issues pertaining to grading and assessment.

- The need for consistency in practice

- The effects zeros have on class averages
- The purpose of summative assessments
- The value and purpose of formative assessments
- The use of grades as indicators of achievement of standards
- The danger of polluting a grade
- The error of ability grouping
- The contribution of grading information systems
- The benefit of high-learning expectations
- The inaccuracy of including effort, behavior, and work ethic within a grade (O'Connor, 2007; Goodwin, 2011; Wormeli, 2009)

Without additional professional development, FRMS will continue to have only pockets of success in best practices with little consistency. With consistent and continuous professional development fully integrated into the school's philosophy and schedule, it will enjoy a full-scale incorporation of best practices in grading and assessment, improving the overall educational experience for every student.

References

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

AMLE. (2010). *This we believe: Keys to educating young adolescents*. Westerville, OH: National Middle School Association.

Annerstedt, C., & Larsson, S. (2010). I have my own picture of what the demands are: Grading in Swedish PEH—problems of validity, comparability and fairness. *European Physical Education Review, 16*(2), 97–115.

Aronson, M. J. (2008). *How teachers' perceptions in the areas of student behavior, attendance and student personalities influence their grading practices* (Doctoral dissertation). Retrieved from ERIC database. (ED3300581)

Arter, J. A., & Chappuis J. (2006). *Creating and recognizing quality rubrics*. Boston, MA: Pearson.

Aseltine, J. M., Faryniarz, J. O., & Rigazio-DiGilio, A. J. (2006). *Supervision for learning.* Alexandria, VA: Association for Supervision and Curriculum Development.

Bailey, K., & Jakicic, C. (2012). *Common formative assessment: A toolkit for professional learning communities at work*. Bloomington, IN: Solution Tree Press.

Balfanz, R. (2009). Putting middle grade students on the graduation path. *Middle School Journal*, 3–15. Retrieved from http://www.amle.org/portals/0/pdf/research /Research_from_the_Field/Policy_Brief_Balfanz.pdf

Barr, R. D., & Parrett, W. H. (2007). *The kids left behind: Catching up the underachieving children of poverty.* Bloomington, IN: Solution Tree Press.

Bartlett, J. C. (2012). Attitudes for excellence. *Principal Leadership, 13*(1), 60–63.

Bonesronning, H. (2004). Do the teachers' grading practices affect student achievement? *Education Economics, 12*(2), 151–167.

Brookhart, S. M. (1994). Teachers' grading: Practice and theory. *Applied Measurement in Education, 7*(4), 279–301.

Brookhart, S. M. (2008). *How to give effective feedback to your students.* Alexandria, VA: Association for Supervision and Curriculum Development.

Brookhart, S. M. (2009). *Grading.* Upper Saddle River, NJ: Pearson.

Brookhart, S. M. (2010). *How to assess higher-order thinking skills in your classroom*. Alexandria, VA: Association for Supervision and Curriculum Development.

Brookhart, S. M. (2011). Starting the conversation about grading. *Educational Leadership, 69*(3), 10–14.

Brookhart, S. M. (2013). Assessing creativity. *Educational Leadership, 70*(5), 28–34.

Brown, D. F., & Knowles, T. (2007). *What every middle school teacher should know.* Westerville, OH: National Middle School Association.

Carey, T. (2001). *An exploratory study of minimum grading practices* (Doctoral dissertation). Retrieved from ERIC database. (ED927604304)

Carey, T., & Carifio, J. (2011, March). Minimum grading, maximum learning. Retrieved from http://www.nassp.org/Content/158/pl_mar11_carifio.pdf

Carlson, L. A. (2003). *Beyond assessment to best grading practice*: *Practical guideline* (Doctoral dissertation). Retrieved from ERIC database. (ED480071)

Chappuis, J. (2009). *Seven strategies of assessment for learning*. Portland, OR: Assessment Training Institute.

Chappuis, J., & Chappuis, S. (2006). *Understanding school assessment.* Boston, MA: Pearson.

Chappuis, J., Stiggins, R., Chappuis, S., & Arter, J. (2012). *Classroom assessment for student learning: Doing it right—using it well.* Upper Saddle River, NJ: Pearson.

Chappuis, S., Chappuis, J., & Stiggins R. (2009). The quest for quality. *Educational Leadership, 67*(3), 14–19.

Chappuis, S., Commodore, C., & Stiggins, R. (2010). *Assessment balance and quality: An action guide for school leaders.* Boston, MA: Pearson.

Copland, M. A., & Knapp, M. S. (2006). *Connecting leadership with learning: A framework for reflection, planning, and action*. Alexandria, VA: Association for Supervision and Curriculum Development.

Dahlgren, L. O., Fejes, A., Abrandt-Dahlgren, M., & Trowald, N. (2009). Grading systems, features of assessment and students' approaches to learning. *Teaching in Higher Education*, *14*(2), 185–194.

Daniels, E. (2010). Creating motivating learning environments: Teachers matter. *Middle School Journal, 43*(2), 32–37.

Danielson, C. (2002). *Enhancing student achievement: A framework for school improvement*. Alexandria, VA: Association for Supervision and Curriculum Development.

Davies, C. A. (2009). *The communicational relationship between assessment types and grading* (Doctoral dissertation). Retrieved from ERIC database. (ED305076619)

DiBattista, D., Gosse, L., Sinnige-Egger, J. A., Candale, B., & Sargeson, K. (2009). Grading schemes, test difficulty, and the immediate feedback assessment techniques. *The Journal of Experimental Education, 77*(4), 311–336.

Doganay, A., & Pinar, A. (2010). The measurement of students' achievement in teaching primary school fifth-year mathematics classes. *Educational Sciences: Theory and Practice, 10*(1), 199–215.

Doubet, K. J. (2012). Formative assessment jump-starts a middle grades differentiation initiative. *Middle School Journal, 43*(3), 32–38.

Dueck, M. (2011). How I broke my own rule and learned to give retests. *Educational Leadership, 69*(3), 72–75.

DuFour, R. (2004). What is a "professional learning community"? *Educational Leadership, 61*(8), 6–11.

DuFour, R., DuFour, R., Eaker, R., & Karhanek, G. (2004). *Whatever it takes: How professional learning communities respond when kids don't learn.* Bloomington, IN: National Educational Service.

DuFour, R., DuFour, R., Eaker, R., & Karhanek, G. (2010). *Raising the bar and closing the gap: Whatever it takes.* Bloomington, IN: Solution Tree Press.

DuFour, R., DuFour, R., Eaker, R., & Many, T. (2010). *Learning by doing*. Bloomington, IN: Solution Tree Press.

DuFour, R., & Marzano R. J. (2011). *Leaders of learning: How district, school, and classroom leaders improved student achievement.* Bloomington, IN: Solution Tree Press.

Duncan, R. C., & Noonan, B. (2007). Factors affecting teachers' grading and assessment practices. *The Alberta Journal of Educational Research, 53*(1), 1–27.

Dweck, C. S. (2006). Mind Sets. *The New Psychology of Success*. New York, NY: Random house.

Dyb, D. C. (2012). *Teacher perceptions of middle level grading practices* (Doctoral dissertation). Retrieved from ERIC database. (ED927932260)

Dyrness, R., & Dyrness, A. (2008). Making the grade in middle school. *Kappa Delta Pi Record, 44*(3), 114–118.

Eaker, R., & Keating, J. (2012). *Every school every team every classroom: District leadership for growing professional learning communities at work*. Bloomington, IN: Solution Tree Press.

Edwards, S. (2013). Preparing new teachers for contemporary middle grades schools. *Middle School Journal*, *44*(4), 14–20.

Elmore, R. (2002). Bridging the gap between standards and achievement: The imperative for professional development in education. *Albert Shanker Institute*, 2–40.

Epstein, J. L. (2005). School-initiated family and community partnerships. In A. Erb (Ed.), *This we believe in action: Implementing successful middle level schools* (pp. 77–96). Westerville, OH: National Middle School Association.

Erickson, J. A. (2001, February). A call to action: Transforming grading practices. *Principal Leadership*, 42–46.

Erickson, J. A. (2006, March). Grading practices: The third rail. *Principal Leadership*, 22–24.

Erickson, J. A. (2011). How grading reform changed our school. *Educational Leadership, 69*(3), 66–70.

Erkens, C., & Twadell, E. (2012). *Leading by design: An action framework for PLC at-work leaders.* Bloomington, IN: Solution Tree Press.

Farrington, C. A. (2008). *Making sense of Fs: How high schools shape students' interpretations of and responses to failure* (Doctoral dissertation). Retrieved from ERIC database. (ED3345175)

Fisher, D., & Frey, N. (2007). *Checking for understanding: Formative assessment techniques for your classroom.* Alexandria, VA: Association for Supervision and Curriculum Development.

Fisher, D., Frey, N., & Pumpian, I. (2011). No penalties for practice. *Educational Leadership, 69*(3), 46–51.

Fleenor, A., Lamb, S., Anton, J., Stinson, T., & Donen, T. (2011). In the grades game. *Principal Leadership, 11*(6). 48–52.

Fullan, M. (2003). *The moral imperative of school leadership*. Thousand Oaks, CA: Corwin Press Inc.

Garrison, C., Chandler, D., & Ehringhaus, M. (2009). *Effective classroom assessment: Linking assessment with instruction.* Westerville, OH: National Middle School Association and Measured Progress.

George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference. 11.0 update* (4th ed.). Boston: Allyn & Bacon.

Gerke, R. J. (2007). *The proper assigning of an F grade: The lack of understanding best practices in grading* (Doctoral dissertation). Retrieved from ERIC database. (ED3379468)

Gliem J. A., & Gliem R. R., (2003). Calculating, interpreting, and reporting Cronbach's Alpha reliability coefficient for Likert-type scales. Retrieved https://scholarworks .iupui.edu/bitstream/handle/1805/344/Gliem%20&%20..?sequence=1

Goodwin, B. (2011). Grade inflation: Killing with kindness? *Educational Leadership, 69*(3), 80–81.

Goodwin, B., & Kirsten, M. (2012). For positive behavior, involve peers. *Educational Leadership, 70*(2), 82–83.

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

Greenstein, L. (2012). Beyond the core: Assessing authentic 21st-century skills. *Principal Leadership, 13*(4), 36–42.

Grimes, T. V. (2010). *Interpreting the meaning of grades: A descriptive analysis of middle school teachers' assessment and grading practices* (Doctoral dissertation). Retrieved from ERIC database. (ED516731)

Gullen, K., Gullen, J., & Erickson-Guy, N. (2012). Conversations about grading. *Principal Leadership*, *12*(8), 24–27.

Gullen, K., & Zimmerman, H. (2013). Saving time with technology. *Educational Leadership*, *70*(6), 63–66.

Guskey, T. R. (2004). Are zeros your ultimate weapon? *Principal Leadership, 14*(5), 31–35.

Guskey, T. R. (2006). *"It wasn't fair!" Educators' recollections of their experiences as students with grading* (Doctoral dissertation). Retrieved from ERIC database. (ED492005)

Guskey, T. R. (2009). *Bound by tradition: Teachers' views of crucial grading and reporting issues* (Doctoral dissertation). Retrieved from ERIC database. (ED509342)

Guskey, T. R. (2011). Five obstacles to grading reform. *Educational Leadership, 69*(3), 16–21.

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

Guskey, T. R., & Jung, L. E. (2012). Four steps in grading reform. *Principal Leadership, 13*(4), 22–28.

Guskey, T. R., Swan, G. M., & Jung, L. A. (2011, October). Grades that mean something. *Kappan Magazine*, *93*(2), 52–57.

Haselhuhn, C. W., Al-Mabuk, R., Gabriele, A., Groen, M., & Galloway, S. (2007). Promoting positive achievement in the middle school: A look at teachers' motivational knowledge, beliefs, and teaching practices. *Research in Middle Level Education*, *30*(9), 1–20.

Hess, F. M. (2010). *The same thing over and over: How school reformers get stuck in yesterday's ideas.* Cambridge: MA. Harvard University Press.

Hummel Wiley, C. R. (2011). *Profiles of teachers grading practices: Integrating teacher beliefs, course criteria, and student characteristics* (Doctoral dissertation). Retrieved from ERIC database. (ED3468465)

IBM SPSS (2013). SPSS statistical software. Retrieved from http://www.01.ibm.com /software/analytics/spss/

Jackson, A. W., & Davis, G. A. (2000). *Turning points 2000: Educating adolescents in the 21st century.* Westerville, OH: National Middle School Association.

James-Ward, C., Frey, N., & Fisher, D. (2012). Root cause analysis: Getting to the root of a problem strengthens school improvement efforts with focused solutions. *Principal Leadership, 13*(2), 59–61.

Johnson, D. (2013). Technology skills every teacher needs. *Educational Leadership, 70*(6), 84–85.

Jung, L. A., & Guskey, T. R. (2011). Fair and accurate grading for exceptional learners. *Principal Leadership, 12*(3), 32–37.

Khan, S., & Slavitt, E. (2013). A bold new math class. *Educational Leadership*, *70*(6), 28–31.

Kinney, P. (2012). Student-led conferences supporting learning. *Principal Leadership, 13*(3), 55–57.

Kinney, P., & Robinson, L. (2005). Courageous, collaborative leadership. In A. Erb (Ed.), *This we believe in action: Implementing successful middle level schools* (pp. 19–27). Westerville, OH: National Middle School Association.

Kohn, A. (2011). The case against grades. *Educational Leadership, 69*(3), 28–33.

Kouzes, J., & Posner, B. (2006). *A leader's legacy.* San Francisco, CA: Jossey-Bass.

Lawrence, J. A. (2011). *An examination of high school assessment and grading practices within a professional learning community* (Doctoral dissertation). Retrieved from ERIC database. (ED3457790)

Lutz, P. A. (2011). What's in your bag of tricks: Differentiating instruction with alternative assessment. *Middle Ground, 15*(1), 12–13.

Mainkar, A. V. (2008). A student-empowered system for measuring and weighing participation in class discussion. *Journal of Management Education, 32*(1), 23–37.

Maloley, K. L. (2008). *Teachers' responses to mandate assessment: Case studies of teachers' assessment practices in elementary classrooms* (Doctoral dissertation). Retrieved from ERIC database. (ED3348152)

Martin, J. (2013). Handing over the keys: Giving student ownership of math. *Middle Ground, 16*(3), 22.

Martinez, J. F., Stecher, B., & Borko, H. (2009). Classroom assessment practices, teacher judgments, and student achievement in mathematics: Evidence from the ESCL. *Educational Assessment, 14*, 78–102.

Marzano, J. R. (2001). *Classroom instruction that works: Research-based strategies for increasing student achievement.* Alexandria, VA: Association for Supervision and Curriculum Development.

Marzano, J. R. (2006). *Classroom assessment and grading that work.* Alexandria, VA: Association of Supervision and Curriculum Development.

Marzano, J. R. (2007). *The art and science of teaching: A comprehensive framework for effective instruction.* Alexandria, VA: Association for Supervision and Curriculum Development.

Marzano, J. R. (2008). *Vision document: Getting serious about school reform, the three critical commitments*. Bloomington, IN: Marzano Research Laboratory.

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

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

Marzano, J. R., Norford, J. S., Paynter, D. E., Pickering, D. J., & Gaddy, B. B. (2001). *A handbook for classroom instruction that works.* Alexandria, VA: Association for Supervision and Curriculum Development.

Marzano, J. R., Pickering, D. J., & Heflebower, T. (2011). *The highly engaged classroom*. Bloomington, IN: Marzano Research Laboratory.

Marzano, J. R., & Waters, T. (2009). *District leadership that works: Striking the right balance.* Bloomington, IN: Solution Tree Press.

Marzano, J. R., Waters, T., & McNulty, A. B. (2005). *School leadership that works: From research to results.* Alexandria, VA: Association for Supervision and Curriculum Development.

Matthews, K. W. (2007). *The relationship among middle level teachers' assessment practices, instructional leadership, and student achievement* (Doctoral dissertation). Retrieved from ERIC database. (ED3322724)

McDaniel, F. (2010). *An investigation of the validity of best grading practices* (Doctoral dissertation). Retrieved from ERIC database. (ED748818306)

McLeod, S. A. (2008). *Simply psychology: Likert scale.* Retrieved from http://www .simplypsychology.org/likert-scale.html

McMillan, J. H., Myran, S., & Workman, D. (2002). Elementary teachers' classroom assessment and grading practices. *Journal of Educational Research, 95*(3),203–213.

McMillan, J. H., & Nash, S. (2000). *Teacher classroom assessment and grading practices decision making* (Doctoral dissertation). Retrieved from ERIC database. (ED447195)

McMillen, Cory (2012). Mirror image: What the accountability debate can teach us about grading. *Association for Middle Level Education's Middle Level Insider***.** Retrieved from http://www.amle.org/Publications/MiddleLeveliNSIDER/Articles/Mirror/tabid/2777/Default.aspx?_cldee=Z3dpbGVzQG5zZDEzMS5vcmc%3d

McMunn, N., Schenck, P., & McColskey, W. (2003). *Standards-based assessment, grading, and reporting in classrooms: Can district training and support change teacher practice?* Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Moore, B. (2012). Transforming your school into a successful learning community. *Middle Ground*, *16*(1), 22–23.

Moss, C. M., & Brookhart, S. M. (2009). *Advancing formative assessment in every classroom: A guide for instructional leaders*. Alexandria, VA: Association of Supervision and Curriculum Development.

Nesin, G. (2005). Students and teachers engaging in active learning. In A. Erb (Ed.), *This we believe in action: Implementing successful middle level schools* (pp. 53–62). Westerville, OH: National Middle School Association.

NCLB. (2013). Retrieved from http://writingcenter.waldenu.edu/640.htm

NMSA. (2010). *This we believe: Keys to educating young adolescents*. Westerville, OH: National Middle School Association.

O'Connor, K. (2007). *A repair kit for grading: 15 fixes for broken grades*. Portland, OR: Educational Testing Service.

O'Connor, K. (2007). The last frontier: Tackling the grading dilemma. In A. Reeves (Ed.), *Ahead of the curve: The power of assessment of transforming teaching and learning* (pp. 127–145). Bloomington, IN: Solution Tree.

O'Connor, K., & Wormeli, R. (2011). Reporting student learning. *Educational Leadership, 69*(3), 41–44.

Patron, H., & Smith, W. J. (2011). Implication of fail–forward in an online environment under alternative grading schemes. *Journal of Educators Online, 8*(1), 1–13.

Popham, W. J. (2008). *Transformative assessment*. Alexandria, VA: Association for Supervision and Curriculum Development.

Popham, W. J. (2011). *Classroom assessment: What teachers need to know*. Boston, MA: Pearson.

Potts, G. (2010). A simple alternative to grading. *Inquiry*, *15*(1), 29–42.

Pridham, B., & Deed, C. (2012). Applied learning and community partnerships improved student engagement in Australia. *Middle School Journal*, *44*(1), 36–42.

Proulx, C., Spencer-May, K., & Westerberg, T. (2012). Moving to standards-based grading: Lessons from Omaha. *Principal Leadership, 13*(4), 30–34.

Ramey, S., VandeVusse, L., & Gosline, M. (2007). Using a written communication rubric to improve students' writing. *International Journal of Learning, 13*(10), 67–74.

Randall, J., & Engelhard, G. (2009). Differences between teachers' grading practices in elementary and middle schools. *The Journal of Educational Research, 102*(3), 175–185.

Reeves, B. D. (2004). The case against the zero. *Phi Delta Kappa, 86*(4), 324–325.

Reeves, B. D. (2007). Introduction. In A. Reeves (Ed.), *Ahead of the curve: The power of assessment of transforming teaching and learning* (pp. 1–12). Bloomington, IN: Solution Tree Press.

Reeves, B. D. (2009). *Leading change in your school: How to conquer myths, build commitment, and get results*. Alexandria, VA: Association for Supervision and Curriculum Development.

Reeves, B. D. (2011). Taking the grading conversation public. *Educational Leadership, 69*(3), 76–79.

Resh, N. (2009). Justice in grades allocation: Teachers' perspective. *Social Psychological Education, 12,* 315–325.

Riddle, M. (2012). School improvement at no cost. *Principal Leadership, 13*(1), 75–77.

Rieg, S. (2007). Classroom assessment strategies: What do students at-risk and teachers perceive as effective and useful? *Journal of Instructional Psychology, 34*(4). Retrieved from http://www.freepatentsonline.com/article/Journal-Instructional-Psychology/17337558.html

Roberts, P. (2006). Reliability and validity in research. *Nursing Standard, 20*(44), 41–45.

Robinson, K. (2012). Teach your children well. *The Costco Connection*, *27*(8), 32–35.

Robinson, K., & Aronica, L. (2009). *The element: How finding your passion changes everything*. New York, NY: Penguin Group.

Rom, C. M. (2011). Grading more accurately. *Journal of Political Science Education, 7* (2), 208–223.

Roorda, N. L. (2008). *The impact of professional development on assessment and grading practices for secondary teachers* (Doctoral dissertation). Retrieved from ERIC database. (ED288264353)

Sadler, R. D. (2010). Fidelity as a precondition for integrity in grading academic achievement. *Assessment and Evaluation in Higher Education,* 727–747. doi:10.1080/02602930902977756

St. Maurice, H., & Yudchitz, A. (2003). *Transforming supervisory grading: Accountability in supervising and evaluating beginning teachers* (Doctoral dissertation). Retrieved from ERIC database. (ED475368)

Salend, J. S. (2011). Creating student-friendly tests. *Educational Leadership, 69*(3), 52–58.

Schenck, J. (2003). *Learning, teaching, and the brain.* Thermopolis, WY: Knowa Inc.

Scherer, M. (2011). What we learn from grades. *Educational Leadership, 69*(3), 7.

Schimmer, T. (2012). *Ten things that matter from assessment to grading*. Toronto, Ontario: Pearson Canada Inc.

Schmoker, M. (1999). *The key to continuous school improvement.* Alexandria, VA: Association for Supervision and Curriculum Development.

Schmoker, M. (2011). *Focus: Elevating the essentials to radically improve student learning.* Alexandria, VA: Association for Supervision and Curriculum Development.

Schurr, S. (2012). *Authentic assessment: Active engaging product and performance measures.* Westerville, OH: Association for Middle Level Education.

Shanahan, K. B. (2011). *Elementary teachers' grading practice: Does the reality reflect the rubric* (Doctoral dissertation)*?* Retrieved from ERIC database. (ED923292642)

Smith, D. (2012). It's about time! *Principal Leadership, 12*(7), 18–22.

Sousa, D. A. (2006). *How the brain learns*. Thousand Oaks, CA: Corwin Press.

Stephens, S. E. (2010). *7th–12th grade English/language arts teachers and their classroom grading practices: Investigating the use of standards-based grading in Nebraska's rural classroom* (Doctoral dissertation). Retrieved from ERIC database. (ED807680640)

Stiggins, R. J. (2001). *Student involvement classroom assessment*. Columbus, OH: Merrill Prentice Hall.

Stiggins, R. J. (2006). Assessment for learning: A key to motivation and achievement. *Phi Delta Kappa International, 2*(2), 3–19.

Stiggins, R. J. (2007). Assessment through the student's eyes. *Educational Leadership, 64*(8), 22–26.

Stiggins, R., Arter, J., Chappuis, J., & Chappuis, S. (2006). *Classroom assessment for student learning*. Columbus, Ohio: Educational Testing Services.

Stiggins, R., & Knight, T. (1998). *But are they learning?* Portland, OR: Assessment Training Institute.

Swaim, S. (2005). A shared vision that guides decisions. In A. Erb (Ed.), *This we believe in action: Implementing successful middle level schools* (pp. 29–34). Westerville, OH: National Middle School Association.

Sylvester, R. (2007). *The adolescent brain: Reaching for autonomy.* Thousand Oaks, CA: Corwin Press.

Tanner, D. (2012). *Using statistics to make educational decisions*. Thousand Oaks, CA: SAGE Publications, Inc.

Teague, G. M., & Anfara, Jr., V. A. (2012). Professional learning communities create sustainable change through collaboration. *Middle School Journal*, *44*(2), 58–63.

Tomlinson, C. A. (2011). Finding your grading compass. *Educational Leadership, 69*(3), 86–87.

Tovani, C. (2012). Feedback is a two-way street. *Educational Leadership, 70*(1), 48–51.

Van de Walle, J. (2001). *Elementary and middle school mathematics*. New York, NY: Addison Wesley Longman.

Varlas, L. (2012). Improving student writing through formative assessments. *Education Update, 54*(2), 5–7.

Vatterott, C. (2011). Making homework central to learning. *Educational Leadership, 69*(3), 60–64.

Villa, R. A., Thousand, J. S., & Nevin, A. I. (2008). *A guide to co-teaching: Practical tips for facilitating student learning.* Thousand Oaks, CA: SAGE Publications, Inc.

Wagner, T. (2002). *Making the grade.* New York, NY: RoutledgeFalmer.

Waters, T., & Cameron, G. (2007). *The balanced leadership: Framework, connecting vision with action.* Denver, CO: Mid-continent Research for Education and Learning.

Wiggins, G., & McTighe, J. (2011). *The understanding by design guide to creating high-quality units.* Alexandria, VA: Association for Supervision and Curriculum Development.

Wilcox, K. C., & Angelis, J. (2009). *Best practices from high-performing middle schools: How successful schools remove obstacles and create pathways to learning*. New York: Teachers College Press.

Wilcox, K. C., & Angelis, J. (2012). From “muddle school” to middle school: Building capacity to collaborate for higher-performing middle schools. *Middle School Journal, 43*(4), 40–48.

Willhoft, J. (2012). Principals leading the way. *Principal Leadership, 13*(4), 18–21.

Willis, J. (2006). *Researched-based strategies to ignite student learning.* Alexandria, VA: Association for Supervision and Curriculum Development.

Winger, T. (2005). Grading to communicate. *Educational Leadership, 63*(3), 61–65.

Wormeli, R. (2006). Accountability: Teaching through assessment and feedback, not grading. *American Secondary Education, 34*(3), 14–27.

Wormeli, R. (2006). *Fair isn’t always equal: Assessing and grading in the differentiated classroom*. Portland, OR: Stenhouse Publisher and National Middle School Association.

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

Wormeli, R. (2012). Making the most of professional development. *Middle Ground, 15*(3), 38–39.

Wormeli, R. (2012). Teaching in the middle: Motivating students to step up. *Middle Ground, 15*(4), 40–41.

Yesbeck, D. M. (2011). *Grading practices: Teachers' considerations of academic and non-academic factors* (Doctoral dissertation). Retrieved from ERIC database. (ED 913076079)

Zmuda, A., Kuklis, R., & Kline, E. (2004). *Transforming schools: Creating a culture of continuous improvement.* Alexandria, VA: Association for Supervision and Curriculum Development.

Zmuda, A. (2010). *Breaking free from myths about teaching and learning.* Alexandria, VA: Association for Supervision and Curriculum Development.

Appendix A

Grading and Assessment Presurvey

Part I

Survey: Grading Practices **Department:_____**

1. By continuing in this survey, you give your consent to participate in this study.
Yes: I do give my consent. No: I do not give my consent.

2. I support the idea of a school-wide grading and assessment policy.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

3. I understand how giving zeros skews the mean (average) in calculating grades.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

4. I believe that grades should principally reflect achievement of standards.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

5. I understand the purpose of formative assessment to be frequent feedback on progress.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

6. It is objective and fair to base grades on a mean of points accumulated in a traditional scale (59% or lower = F; 60–69% = D; 70%–79% = C; 80–89% = B; 90–100% = A).
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

7. It is okay to give a student an A if he or she has mastered the course objective but not completed all homework and class work.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

8. I believe students who turn in late work should be penalized for missing deadlines.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

9. I think teachers should be able to adopt whatever grading practices work for them.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

10. Extra credit should not be allowed in any class.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

11. Using a grade of incomplete is preferable to penalizing late work or giving zeros or Fs.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

Appendix A

Grading and Assessment Presurvey

Part II

12. I believe grades should reflect the influence of a student's behavior and work ethic on his or her achievement.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

13. Grades are comparative in nature and serve to sort students into ability groups.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

14. I understand how to develop high-quality summative assessments.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

15. I understand the purpose of summative assessment to be a periodic report of achievement of standards.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

16. In figuring summative grades, recent evidence is more valuable than an average of points over time.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

17. My students are routinely involved in formative assessment of their learning.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

18. The grading information system (e.g., PowerSchool) has interfered with my grading practice.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

19. High expectations for learning are communicated in the overall school setting.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

20. Teachers should enter grades at least once a week into the grading information system (e.g., PowerSchool).
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

Appendix A

Grading and Assessment Presurvey

Part III

21. It is okay to lower a student's grade based on disruptive behavior.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

22. If we did not grade student work, they would not do it.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

23. Behavior and effort are important factors in academic achievement and should be reflected in student grades.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

Appendix B

Grading and Assessment Postsurvey

Part I

Survey: Grading Practices **Department:_____**

Did you take the presurvey in the Fall? Circle one: YES NO

1. By continuing in this survey, you give your consent to participate in this study.
Yes: I do give my consent. No: I do not give my consent.

2. I support the idea of a school-wide grading and assessment policy.
Strongly Agree Agree Neutral Disagree Strongly Disagree

3. I understand how giving zeros skews the mean (average) in calculating grades.
Strongly Agree Agree Neutral Disagree Strongly Disagree

4. I believe that grades should principally reflect achievement of standards.
Strongly Agree Agree Neutral Disagree Strongly Disagree

5. I understand the purpose of formative assessment to be frequent feedback on progress.
Strongly Agree Agree Neutral Disagree Strongly Disagree

6. It is objective and fair to base grades on a mean of points accumulated in a traditional scale (59% or lower = F; 60–69% = D; 70%–79% = C; 80–89% = B; 90–100% = A).
Strongly Agree Agree Neutral Disagree Strongly Disagree

7. It is okay to give a student an A if he or she has mastered course objective but not completed all homework and class work.
Strongly Agree Agree Neutral Disagree Strongly Disagree

8. I believe students who turn in late work should be penalized for missing deadlines.
Strongly Agree Agree Neutral Disagree Strongly Disagree

9. I think teachers should be able to adopt whatever grading practices work for them.
Strongly Agree Agree Neutral Disagree Strongly Disagree

10. Extra credit should not be allowed in any class.
Strongly Agree Agree Neutral Disagree Strongly Disagree

11. Using a grade of incomplete is preferable to penalizing late work or giving zeros or Fs.
Strongly Agree Agree Neutral Disagree Strongly Disagree

Appendix B

Grading and Assessment Postsurvey

Part II

12. I believe grades should reflect the influence of a student's behavior and work ethic on his or her achievement.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

13. Grades are comparative in nature and serve to sort students into ability groups.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

14. I understand how to develop high-quality summative assessments.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

15. I understand the purpose of summative assessment to be a periodic report of achievement of standards.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

16. In figuring summative grades, recent evidence is more valuable than an average of points over time.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

17. My students are routinely involved in formative assessment of their learning.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

18. The grading information system (e.g., PowerSchool) has interfered with my grading practice.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

19. High expectations for learning are communicated in the overall school setting.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

20. Teachers should enter grades at least once a week into the grading information system (e.g., PowerSchool).
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

Appendix B

Grading and Assessment Postsurvey

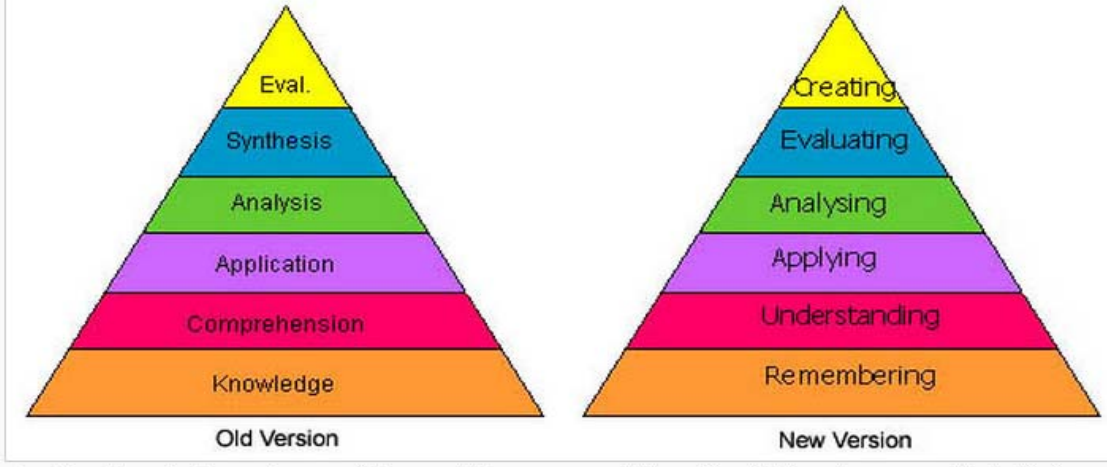
Part III

21. It is okay to lower a student grade based on disruptive behavior.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

22. If we did not grade student work, they would not do it.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

23. Behavior and effort are important factors in academic achievement and should be reflected in student grades.
Strongly Agree **Agree** **Neutral** **Disagree** **Strongly Disagree**

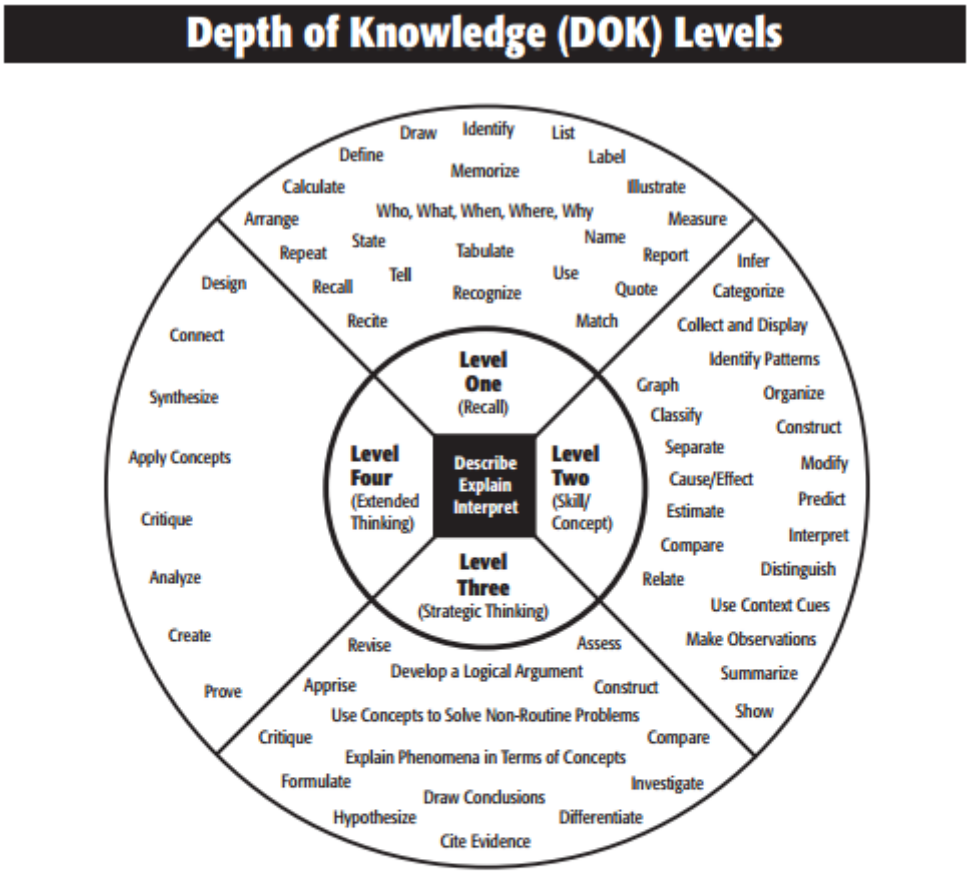
Appendix C



Caption: Terminology changes *"The graphic is a representation of the NEW verbage associated with the long familiar Bloom's Taxonomy. Note the change from Nouns to Verbs [e.g., Application to Applying] to describe the different levels of the taxonomy. Note that the top two levels are essentially exchanged from the Old to the New version." (Schultz, 2005) (Evaluation moved from the top to Evaluating in the second from the top, Synthesis moved from second on top to the top as Creating.) Source: http://www.odu.edu/educ/roverbau/Bloom/blooms_taxonomy.htm*

Figure 23. Old and New Bloom's Taxonomy. Source: http//www.odu.edu/educ/roverbau /Bloom/blooms_taxonomy.htm (2013)

Appendix D



Level One Activities	Level Two Activities	Level Three Activities	Level Four Activities
Recall elements and details of story structure, such as sequence of events, character, plot and setting.	Identify and summarize the major events in a narrative.	Support ideas with details and examples.	Conduct a project that requires specifying a problem, designing and conducting an experiment, analyzing its data, and reporting results/ solutions.
Conduct basic mathematical calculations.	Use context cues to identify the meaning of unfamiliar words.	Use voice appropriate to the purpose and audience.	Apply mathematical model to illuminate a problem or situation.
Label locations on a map.	Solve routine multiple-step problems.	Identify research questions and design investigations for a scientific problem.	Analyze and synthesize information from multiple sources.
Represent in words or diagrams a scientific concept or relationship.	Describe the cause/effect of a particular event.	Develop a scientific model for a complex situation.	Describe and illustrate how common themes are found across texts from different cultures.
Perform routine procedures like measuring length or using punctuation marks correctly.	Identify patterns in events or behavior.	Determine the author's purpose and describe how it affects the interpretation of a reading selection.	Design a mathematical model to inform and solve a practical or abstract situation.
Describe the features of a place or people.	Formulate a routine problem given data and conditions.	Apply a concept in other contexts.	
	Organize, represent and interpret data.		

Figure 24. *Depth of Knowledge*. Source:

http://www.missouristate.edu/assessment/90018 .htm (2013)

Appendix E

Conversion of Scale Score to Percentage

Label	Scale Score	% on 10-point task
Advanced	4	10
Proficient	3	8–9
Basic	2	7
Below Basic	1	6
Missing or Incomplete	M or I	5

	Label	Scale Score	% on 25-point task
A	**Advanced**	4	23–25
B	**Proficient**	3	20–22
C	**Basic**	2	17–19
D	**Below Basic**	1	15–16
F	**Missing or Incomplete**	M or I	12.5

Appendix F

Self-Assessment— Eighth-Grade Science

Instructions: Look at the four topics that will be covered on the test. Highlight the information that you understand. Then in the areas that you find you are not Proficient in, please practice by completing the practice problems.

Metrics

Below Basic:
- Volume measured in Liters
- Length measured in Meters
- Mass measured in Grams

Basic:
- You can read a ruler, beaker, cylinder, scale.

Proficient
- You can calculate density and know the units of density.

Advanced
- You can convert between metric units.
  3.2 m = _____km

PRACTICE:
Do the following problems from the book:
Pg. 14 questions 1 (all), 3 (all)
Pg. 31 questions 7
Pg. 45 question 2 (all)
Pg. 67 questions 2, 7

The Air Around You

Below Basic:
- You can define atmosphere.
- You know that the atmosphere provides all the gases necessary for life on Earth.

Basic:
- You know that nitrogen makes up most of the atmosphere.
- You know that oxygen makes up 21% of the atmosphere.

Proficient:
- You can explain why you should shut all the windows in your house when there is a fire.
- You can explain why our lab shows the amount of oxygen in the air.

Advanced:
- You can calculate the amount of oxygen in the air using the equation from the lab.

PRACTICE:
Do the following problems from the book:
Pg. 9 question 1b
Pg. 21 questions 1–2 (all letters)
Pg. 31 questions 1, 3, 4, 8, 9

Appendix F (cont.)

Density

Below Basic
- What is density?

Basic
- How do you calculate density?

Proficient
- You are able to calculate density.
- As air pressure increases, density _____.
- As altitude increases, density _____.
- Know that warm air rises!

Advanced
- If given a story problem, you can pull out the necessary information and decide what material is more or less dense.

PRACTICE:

Do the following problems from the book:

Pg. 14 questions 1–2 (all)
Calculate density if:
I have a mass of 1 and a volume of 2
I have a volume of 5 and a mass 3
I have a mass of 5 and a volume of 7
I have a volume of 6 and a mass of 9

Air Pressure

Below Basic
- Be able to define air pressure.

Basic
- Know that as altitude increases, air pressure decreases.

Proficient
- Be able to describe why an experiment/demonstration shows air pressure.
- Be able to describe why density and air pressure are so closely related.

Advanced
- Be able to explain weather phenomenon that occur because of air pressure

PRACTICE:

Do the following problems from the book:

Pg. 33 questions 3–5
Pg. 39 questions 1–3 (all)

Appendix G

Astronomy Test—Historical Discoveries

Name_____

Date_____ Period_____

*Answer **EVERY** question to the best of your ability. Even if you have to guess WRITE SOMETHING, for EVERY QUESTION!! If you answer EVERY QUESTION, even if you get it wrong, the lowest score you can get on this test is a 50%. If you need additional room to write, you may attach a lined sheet of paper to the test.*

Standard 1.2.1

1. (BB—Remember/Understand) Astronomy is the study of_____.
2. (BB—Remember/Understand) Recorded history of astronomy dates back as far as
    a. 1609 Galileo
    b. 3000 BC Egyptians
    c. 1969 Neil Armstrong
    d. 1543 Copernicus
3. (B) List 3 historical events that have occurred in the history of astronomy. Note: I am not looking for dates.

    _____
    _____
    _____
4. (P) Summarize the history of astronomy from the first recorded events to current happenings in 3–5 sentences.

    _____
    _____
    _____
    _____
    _____
    _____
    _____
5. (A) Contrast how the study of astronomy has changed throughout history. Use 3 historical events to support your explanation; they can be the same events as you listed in number 4 or different.

    _____
    _____
    _____
    _____
    _____

Appendix G (cont.)

Standard 1.1.2

6. (BB) Draw a diagram of our solar system. HELP: The planets in order are Mercury, Venus, Earth, Mars, Jupiter, Saturn, Uranus, Neptune, (and dwarf planet Pluto if you want to include him).
7. (B) Add to your drawing in number 6 the planets' orbits—very important that you draw them as ellipses.
8. (B) Label your drawing in number 6 as heliocentric or geocentric.
9. (P) Explain one way that Galileo was able to convince the people that the Earth revolves around the Sun.

   _____

   _____
10. (P) Explain a second way that Galileo was able to convince the people that the Earth revolves around the Sun.

    _____

    _____
11. (P-A) Draw a picture explaining what you wrote for number 9.
12. (P-A) Draw a picture explaining what you wrote for number 10.
13. (A) Consider that the Sun and Moon rise, move across the sky, and set every day. If the Earth is not the center of the solar system with everything moving around it, why do the Sun and Moon appear to be going around the Earth?

    _____

    _____

Standard 1.2.2

14. (BB) Draw a refracting telescope.
15. (B) Label any lenses or mirrors as "lens" or "mirror" in your question 14 drawing.
16. (P) Label any lenses or mirrors as concave or convex in your question 14 drawing.
17. (A) Explain how a refracting telescope works. Include focal point in your explanation.

    _____

    _____

    _____
18. (BB) Draw a reflecting telescope.
19. (B) Label any lenses or mirrors as "lens" or "mirror" in your question 18 drawing.
20. (P) Label any lenses or mirrors as concave or convex in your question 18 drawing.

Appendix H

Reviewing My Math Results

Part I

Name _____ Assignment _____ Date ____

Please look at your corrected test or assignment and mark whether each problem is Right or Wrong. Then look at the problems you got wrong, and decide if you made a simple mistake. If you did, mark the "Simple Mistake" column. For all the remaining problems you get wrong, mark the "Don't Get It" column.

Problem	Learning Target	Right	Wrong	Simple Mistake	Don't Get It

Part II

Do I need to retake this assessment?
If you answered yes, please complete the necessary practice problems.

Problems missed	Recommended practice (pages and problems)	Yes	No
1–3			
4–6			
7–10			

Part III

Reflection:

What did you do to prepare for this quiz (formative assessment) and what helped you the most?

How do you plan to study for the summative chapter ___Test?

Appendix I

Example Math Game: Seventh Grade

Basketball player "A" made 23 out of 34 free throws over the course of three games.

- What is his free throw percentage?

A professional basketball court is 94 ft. long and 50 ft. wide.

- What is the total area of the court?

If a player had to run a perimeter, how far would they have to go?

The point guard scored 14 points in the last basketball game. Baskets are worth, 1, 2, and 3 points.

- Show five different ways he could have scored his 14 points.

Last year our Basketball Coach had a win/loss ratio of 3:2. If he had that exact same ratio this year, and the team plays 50 games, how many games will the coach win this year?

At halftime, the team's organization does a raffle that raises $750—25% goes to the professional team's organization, 50% goes to the Boys and Girls Club, and 25% goes to the winning fan.

- How much money does the winning fan receive?

The teams starting forward stands 6 feet 6 inches tall. How many inches is this?

Monday	30 min.
Tuesday	22 min.
Wednesday	45 min.
Thursday	15 min.
Friday	26 min.
Saturday	30 min.
Sunday	? min.

Based off of the table, what is the mean practice time for shooting?

Median?

Mode?

If the Coach wants the teams players to practice 30 minutes a day on shooting, how many minutes would they have to practice on Sunday to have a mean time of 30 minutes?

Appendix J

Levels of Achievement

Name: _____ 7th Grade Math Chapter 2: Integers and Rational Numbers **Assignments/Quizzes/Tests**	My scorepoints possible	1.1.1 Compare magnitudes of integers.	1.1.2 Solve problems requiring the conversion between simple decimals, fractions, ratios, and percents.	1.1.3 Locate the position of rational numbers on a number line.	1.1.5 Apply the number theory concepts of primes, composites, and prime factorization and find the LCM and GCF.	1.2.2 Add, subtract, multiply and divide integers.	3.1.3 Use symbols to express relationships.	3.3.1 Solve one-step equations.	3.1.1 Use variables in simple expressions and equations.	3.1.2 Translate word statements into algebraic expressions & equations.	Use appropriate vocabulary and symbols.			
Pre-Quiz 2.1 – 2.5														
Weekly Homework #5														
Recommended Practice #5														
Quiz 2.1 – 2.5														
Pre-Quiz 2.9 – 2.11														
Weekly Homework #6														
Recommended Practice #6														
Quiz 2.9 – 2.11														
Pre-Quiz 2.6 – 2.8														
Weekly Homework #7														
Recommended Practice #7														
Quiz 2.6 – 2.8														
Chapter 2 Test														

Chapter Rubric
4: All problems correct
3: Most problems correct
2: Some problems correct
1: Minimal problems correct

Appendix K

Learning Targets

Name _____

Mark whether each problem is right or wrong. Then look at the problems you got wrong and decide if you made a simple mistake If you did, mark the "simple mistake" column. For all the remaining problems you got wrong, mark the "Don't get it" column.

Quiz

Problem	Learning target	Right	Wrong	Simple Mistake	Don't Get It
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					

Do I need to re-take this assessment? yes no

If you answered yes; please complete the necessary practice problems

Problems missed	Recommended Practice (pages and problems)		
1-3			
4-6			
7-10			

Reflection

What did you do to prepare for this quiz and what helped you the most?

How do you plan to study for the summative chapter _ test?

Appendix L

Formative Versus Summative

Formative	**Summative**
Notes	Tests (Proficiency Based)
Warm-ups	Projects (Rubric Scored)
Homework	
Worksheets	
Quizzes	
Projects	
20% or less of student grade	80% or more of students grade

Appendix M

Proficiency Scoring

Scoring Scale (Marzano)

This scale helps to distinguish between higher and lower level cognitive activities.

Score	Description of Place on Scale
4.0	**In addition to Score 3.0 performance, in-depth inferences and applications that go beyond what was taught**
3.5	In addition to Score 3.0 performance, partial success at inferences and applications that go beyond what was taught
3.0	**No major errors or omissions regarding any of the information and/or processes (simple or complex) that were explicitly taught**
2.5	No major errors or omissions regarding the simpler details and process and partial knowledge of the more complex ideas and processes
2.0	**No major errors or omissions regarding the simpler details and processes, but major errors or omissions regarding the more complex ideas and processes**
1.5	Partial knowledge of the simple details and processes but major errors or omissions regarding the more complex ideas and procedures
1.0	**With help, a partial understanding of some of the simpler details and processes and some of the more complex ideas and processes**
0.5	With help, a partial understanding of some of the simpler details and processes but not the more complex ideas and processes
0.0	**Even with help, no understanding or skill demonstrated**

Proficiency grading

Scale	Description
4	Advanced / Master
3	Proficient
2	Basic / Apprentice
1	Below Basic / Novice

Appendix N

Learning Strands

Name: _____ 7th Grade Math Chapter 2: Integers and Rational Numbers **Assignments/Quizzes/Tests**	**My score/points possible**	**1.1.1 Compare magnitudes of integers.**	**1.1.2 Solve problems requiring the conversion between simple decimals, fractions, ratios, and percents.**	**1.1.3 Locate the position of rational numbers on a number line.**	**1.1.5 Apply the number theory concepts of primes, composites, and prime factorization and find the LCM and GCF.**	**1.2.2 Add, subtract, multiply and divide integers.**	**3.1.3 Use symbols to express relationships.**	**3.3.1 Solve one-step equations.**	**3.1.1 Use variables in simple expressions and equations.**	**3.1.2 Translate word statements into algebraic expressions & equations.**	**Use appropriate vocabulary and symbols.**			
Pre-Quiz 2.1 – 2.5														
Weekly Homework #5														
Recommended Practice #5														
Quiz 2.1 – 2.5														
Pre-Quiz 2.9 – 2.11														
Weekly Homework #6														
Recommended Practice #6														
Quiz 2.9 – 2.11														
Pre-Quiz 2.6 – 2.8														
Weekly Homework #7														
Recommended Practice #7														
Quiz 2.6 – 2.8														
Chapter 2 Test														

Chapter Rubric
4: All problems correct
3: Most problems correct
2: Some problems correct
1: Minimal problems correct

Appendix O

Homework Completion

6th Grade - Home Work Completion Program					Teacher Directed	Teacher Directed	Administration
Check In:	Name:	Assignment(s)	Assignments(s)	Assignments(s)	Opportunity 1 (Teacher)	Opportunity 2 (Teacher)	Not an Option
	Student A - Example	Reading Response					
	Student B - Example	Spelling 3-1	RR10				
		Spelling 3-1					
		pg. 240-241 (1-14 and 50-55					
		pg. 240-241 (1-14 and 50-55	RR10				
		pg. 230-231 (1-14 and 41-52	pg 236-237 (1-7 and 48-53)				
		pg. 240-241 (1-14 and 50-55	pg. 246 (16-30)	RR10			
		pg. 230-231 (1-14 and 41-52	pg. 236 (16-25)				
		pg. 236 (16-25)	pg. 230-231 (1-14 and 41-52)				
		pg. 236 (16-25)	pg. 230-231 (1-14 and 41-52)				

Appendix P

Grading and Assessment Policy

The intent of this policy is to provide congruence and consistency between and among departments and grade levels in the use of grading and assessment procedures. Since one of the fundamental roles of grades is to provide information to students and parents about student progress, we include their roles and responsibilities in this emerging policy as well.

The staff at ??? School believes…

1. Assessment should be frequent and provide meaningful feedback.
   - Teachers will provide students with ongoing and descriptive feedback on their learning to help them establish goals for improvement.
   - Learning expectations and criteria for assessment are communicated to students in advance.
   - Students are provided with opportunities to learn how to assess their own work and to set goals for improvement.
2. Grades should reflect achievement of standards.
   - Teachers work collaboratively to determine achievement levels and to establish exemplars.
3. Student behavior (Respect and Relationships) and engagement (Readiness and Responsibility) will be assessed separately from the academic grades based on the rubrics created by ??? staff. A student's behavior in the classroom and engagement in his/her learning is vital to intellectual and skill development. Work ethic and the ability to get along with others are life-long skills that transfer to every career path. Evaluations for behavior and engagement are taken very seriously.
4. Cheating, Plagiarism, Academic Dishonesty are all clearly defined in the student handbook. At ??? School, students <u>will not</u> be given a zero for work that is determined to be fraudulent. They will be expected to redo and complete the work honestly. Punishment for cheating will be handled as a behavior infraction; consequences will be determined depending on the severity of the academic dishonesty.
5. Determination of grading levels for formal reporting purposes should primarily reflect student performance on summative tasks. Students' grades will reflect their most consistent and recent level of achievement at the time of reporting.
   - **Summative Assessments** represent about 80% of a student's grade. They may include tests, projects, writing tasks, reflections on simulations, or lab assessments.
   - **Formative Assessments** represent about 20% of a student's grade. They may include homework, class activities, practice, rough drafts, lab activities, and quizzes.

Appendix P (Cont.)

- It is important to note that although formative assessments represent only 20% of the final student grade, they represent the essential activities and practices that stimulate the learning and the primary investment of teacher and student time. Without the practice, the drill, the trials, the drafts—done in a low-risk, supportive environment—the students would not be able to meet the challenges of summative assessments meant to gauge the students' skills against standards. A sports metaphor is most apt; a student practices regularly under the guidance of a skilled coach in order to meet the challenge of the game in which one puts his skills on the line.

6. Teachers avoid grading practices that distort the meaning of a grade by underinflation or overinflation.
    - Extra credit is not allowed when it attempts to fill in missing scores with tasks unrelated to the work required by standards.
    - We discourage reducing the value of school work turned in late. All work should be scored against standards. Missing deadlines, however, is a bad habit that should be reported in the Behavior and Engagement grade. When students repeatedly turn in work late without reasonable explanations, they put themselves in jeopardy of quickly falling behind.
        - Progress reports (at midquarter) serve as an important reminder of grade standing, progress, and work completion.
        - About two weeks before the end of each quarter, the school posts Firm Deadlines, after which no more late or missing work may be turned in.
    - The use of zeros to mark missing work skews students' grades dramatically downward. Instead of using zeros, teachers are strongly encouraged to hold students accountable for completing the missing work and use other codes listed below to communicate a student's status.
        - M indicates work is missing.
        - I indicates the work is incomplete or not yet done to standards. When the teacher *has no evidence of student learning because of significant missing work*, the teacher will mark the *course grade* as "I" and with no credit value (0%). The purpose of the mark is to alert both parent and student that the student is in jeopardy of failing the class.
7. Making the change from assigning points to marking proficiency level is recommended practice. Examples of a rubric defining proficiency levels and conversion charts for changing scale scores to percents can be found in the addendum to this policy.
8. Departments may have more specific assessment policies and will summarize those in letters to parents at the beginning of the school year. All teachers will send a statement of assessment and grading policy to parents at the beginning of the year.

Appendix P (Cont.)

Students

??? School places an unrelenting emphasis on **student responsibility** for learning. To that end we expect

- All work must be done on time.
- All work should meet standards.
- Students should know and plan for firm deadlines nearing the end of marking periods.
- Students should use formative assessments to identify strengths and weaknesses and, when needed, seek extra help.
- Capable students who intentionally do not complete course work should expect immediate and natural consequences in the form of after-school and/or lunch-time detentions, and phone calls home.

Parents

- May confer with a teacher about student progress any time during the year
- Have online access to student progress every day of the week, 24 hours a day, through PowerSchool
- Should understand that the PowerSchool “gradebook” is a record-keeping device that intends to report progress toward achievement of standards, and as such, changes frequently as indicators of growth are evaluated. No reporting devices, however, can replace the power and effect of communications between teachers, students, and parents.
- Should advise their students to seek extra help when formative assessment indicates the student is struggling
- Will receive formal progress indicators eight times a year: progress reports at each midquarter and quarter grades every nine weeks.
- Progress reports are formative in nature and provide feedback to students and parents about strengths and weaknesses.
- Report cards are summative in nature and capture a picture of achievement of standards after nine weeks of instruction.

Appendix Q

Single-Skill Rubric

6th Grade Writing Rubric – Research Writing

Organization

W.6.2 Write informative/explanatory texts to examine and convey complex ideas, concepts, and information through the selection, organization, and analysis of relevant content.

W.6.4 Produce clear and coherent writing in which the development, organization, and style are appropriate to task, purpose, and audience

Advanced	Proficient	Basic	Below Basic
• Writing is developed with several relevant facts, definitions, details, quotes, and/or other examples. • Writing is organized to include: introduction, body paragraphs, and a concluding statement. • Writing utilizes transitions to purposefully move between different pieces of information. • Writing uses a formal style relevant to the audience.	• Writing is developed with relevant facts, definitions, details, quotes, and/or other examples. • Writing is organized to include: introduction, body paragraphs, and a concluding statement. • Writing utilizes transitions between different pieces of information • Writing uses a formal style relevant to the audience.	• Writing is developed with few facts, definitions, details, quotes, or other examples. • An attempt is made to organize the writing. • Transitions between different pieces of information are missing or are ineffective. • Writing does not show an awareness of audience.	• Writing may be a list of facts, have insufficient detail, or may be a summary. • Writing is disorganized and is missing structure. • Transitions are missing • Writing does not show an awareness of audience.

Ideas/Content

W.6.8 Gather relevant information from multiple print and digital resources, assess the credibility and accuracy of each source, and integrate the information while avoiding plagiarism

W.6.9 Draw evidence from literary or informational texts to support analysis, reflection, and research

Advanced	Proficient	Basic	Below Basic
• Writer uses multiple credible sources. • Writer avoids plagiarism • Writer correctly quotes, paraphrases, and summarizes information. • Writer correctly cites sources using a bibliography or a works cited page.	• Writer uses credible sources. • Writer avoids plagiarism • Writer quotes, paraphrases, and summarizes information. • Writer cites sources using a bibliography or a works cited page.	• Writer uses at least one source. • Writer makes an attempt to quote, paraphrase, or summarize information. • Writer makes an attempt to cite sources.	• Use of source information is not evident or is unrelated to the topic. • Writer does not quote, paraphrase, or summarize information. • Writer does not cite sources.

Sentence Fluency/Conventions

L.6.2 Demonstrate command of the conventions of standard English capitalization, punctuation, and spelling when writing.

Advanced	Proficient	Basic	Below Basic
• There are very few capitalization, punctuation, and spelling, and they do not distract from the readability.	• There are some capitalization, punctuation, and spelling errors, but they do not distract from the readability.	• There are several capitalization, punctuation, and spelling, errors which may distract from the readability.	• There are multiple capitalization, punctuation, and spelling errors which make understanding difficult.

Appendix R

Multiple-Skill Rubric

(PRESENTATION COMPONENT) **7th Grade Writing Rubric – Research Writing**

SL.7.4 Present claims and findings, emphasizing salient points in a focused, coherent manner with pertinent descriptions, facts, details, and examples; use appropriate eye contact, adequate volume, and clear pronunciation.			
Advanced • Information is presented in a logical sequence. • Presenter uses appropriate tone, volume, eye contact, and pronunciation. • Presenter focuses on salient points/information.	**Proficient** • Information is presented in a logical sequence. • Presenter uses appropriate tone, volume, eye contact, and pronunciation. • Presenter mostly focuses on salient points/information.	**Basic** • Information presented has some organizational structure. • Presenter may use appropriate tone, volume, eye contact, and pronunciation. • Presenter mostly focuses on pertinent information.	**Below Basic** • Information presented is unstructured. • Presenter seldom uses appropriate tone, volume, eye contact, and pronunciation. • Presentation is lacking information.

SL.7.5 Include multimedia components and visual displays in presentations to clarify claims and findings and emphasize salient points.			
Advanced • Media or visual display enhances information in a meaningful manner.	**Proficient** • Media or visual display supports presented information.	**Basic** • Media or visual display is present, but may be unrelated or ineffective.	**Below Basic** • Media or visual display is unrelated to the topic or is missing.

Appendix S

Kid-Friendly Rubric

6th Grade Writing Rubric – Research Writing

Organization

W.6.2 Write informative/explanatory texts to examine and convey complex ideas, concepts, and information through the selection, organization, and analysis of relevant content.

W.6.4 Produce clear and coherent writing in which the development, organization, and style are appropriate to task, purpose, and audience

Advanced	Proficient	Basic	Below Basic
• Lots of facts, definitions, details, quotes, and/or other examples. • Uses an introduction, body paragraphs, and conclusion. • Transitions are strong and smoothly move the reader from one paragraph to the next. • The author's voice/language are appropriate for the audience.	• Some facts, definitions, details, quotes, and/or other examples. • Uses an introduction, body paragraphs, and conclusion. • Transitions are evident, but may be clunky. • The author's voice/language are appropriate for the audience.	• Just a list of facts, • Some paragraphing; a bit disorganized. • Transitions are unclear or are missing. • Writing shows use of slang or other inappropriate language.	• Information is not relevant to the topic. • Very disorganized. • No transitions. • Writing shows use of slang or other inappropriate language.

Ideas/Content

W.6.8 Gather relevant information from multiple print and digital resources, assess the credibility and accuracy of each source, and integrate the information while avoiding plagiarism

W.6.9 Draw evidence from literary or informational texts to support analysis, reflection, and research

Advanced	Proficient	Basic	Below Basic
• Uses many credible sources. • You don't plagiarize. • You correctly quote, paraphrases, and summarize information. • You correctly cite your sources with a bibliography or a works cited page.	• You uses at least one credible source. • You try to avoid plagiarism • You quote or paraphrase or summarize your information. • You cite your sources using a bibliography or a works cited page.	• You use at least one source, but it might not be a good source. • You try to quote, paraphrase, or summarize information, but it is done incorrectly. • You try to cite sources, but it is done incorrectly.	• You did not use any other sources of information, or • You did not cite any other sources. • You plagiarized your work

Sentence Fluency/Conventions

L.6.2 Demonstrate command of the conventions of standard English capitalization, punctuation, and spelling when writing.

Advanced	Proficient	Basic	Below Basic
• No capitalization, punctuation, or spelling errors.	• Some capitalization, punctuation, and spelling errors, but they do not distract from the readability.	• The capitalization, punctuation, and spelling, errors may distract from the readability.	• Many capitalization, punctuation, and spelling errors that make understanding difficult.

Appendix T

Professional Development: Grading and Assessment Meeting #1

Meeting #1
October 31, 2012

Prework: Read chapter 1 in *A Repair Kit for Grading* by Ken O'Connor
Read journal article, "What We Learn from Grades," by Marge Scherer
Read journal article, "Starting the Conversation about Grading," by Susan M. Brookhart

Agenda:

1. **Read sections of the Preface to the whole group**

2. **Review the reading:** Divide into Vertical PLCs and discuss each bulleted item.

   Chapter 1: "Setting the Stage" Grades must be…
   - Consistent
   - Accurate
   - Meaningful
   - Supportive of learning

3. **Select one person per Vertical PLC to report out to the whole group**

Appendix U

Professional Development: Grading and Assessment Meeting #2

Meeting #2
November 7, 2012

Prework: Read chapter 2 in *A Repair Kit for Grading* by Ken O'Connor
Read chapter 8 in *Fair Isn't Always Equal* by Rick Wormeli
Look at example rubric for Behavior and Engagement (see Appendix M).

1. **Review the reading:** Divide into Vertical PLCs and discuss each bulleted item.

   Chapter 2: "Fixes for Practices that Distort Achievement"

   Don't include…
   - Student behavior
   - Late work penalty
   - Points for extra credit
   - Penalty for cheating
   - Points for attendance
   - Group scores

2. **Please answer the Key Question:** How to implement these "fixes" and still hold students accountable for being responsible for behavior and engagement?

3. **Select one person per Vertical PLC to report out to the whole group concerning the discussions surrounding the key question.**

Appendix V

Professional Development: Grading and Assessment Meeting #3

Meeting #3
November 21, 2012

Prework: Read chapter 3 in *A Repair Kit for Grading* by Ken O'Connor
Read journal article, "Five Obstacles to Grading Reform," by Thomas R. Guskey
Read journal article, "How I Broke My own Rule and Learned to Give Retests," by Myron Dueck

1. **Review the reading:** Divide into Vertical PLCs and discuss each bulleted item.

   - Fix # 7: Organizing grades by standards
     - Show examples
   - Fix # 8: Provide clear descriptions of achievement expectations
     - Using rubrics with clearly described levels of achievement and expectation
   - Fix #9: Compare each student's performance to preset standards
     - Look at Common Core
   - Fix #10: Rely only on quality assessments
     - Show examples

2. **Key Question**: How can I use my grades to reflect achievement of standards?

3. **Select one person per Vertical PLC to report out to the whole group**

Appendix W

Professional Development: Grading and Assessment Meeting #4

Meeting #4
December 5, 2012

Prework: Read chapter 4 in *A Repair Kit for Grading* by Ken O'Connor
Read chapter 9 in *Fair Isn't Always Equal* by Rick Wormeli
Read—Summary of the Power of I (see Appendix CC)

1. **Review the reading:** Divide into Vertical PLCs and discuss each bulleted item.

   - Consider other measures of central tendency and use professional judgment.

   - Use alternatives, such as reassessing to determine real achievement, or use "I" for Incomplete or Insufficient evidence.

2. **Key Question:** Will eliminating zeros for missing work send the wrong message to kids, making them think the work is not important?

3. **Select one person per Vertical PLC to report out to the whole group concerning the discussions surrounding the key question.**

Appendix X

Professional Development: Grading and Assessment Meeting #5

Meeting #5
December 19, 2012

Prework: Read chapter 5 in *A Repair Kit for Grading* by Ken O'Connor
Read journal article, "Reporting Student Learning," by Ken O'Conner and Rick Wormeli
View—Sample Grading and Assessment School-Wide Policy (see Appendix P)

1. **Review the reading:** Divide into Vertical PLCs and discuss each bulleted item.

   - Use only summative evidence
   - Emphasize more recent achievement
   - Student roles

2. **Key Questions:**

   - How will our teaching practice change if we believe grades should reflect achievement on standards?
   - Do we need a policy to help our grading and assessment practices?

3. **Select one person per Vertical PLC to report out to the whole group concerning the discussions surrounding the key question.**

Appendix Y

Teacher Experience

Fall River Middle School
Years of Experience in Teaching
2012–2013

Last	First	Years Teaching	Highest Degree earned	Masters Yes or No	Male	Female
		26	BA	N	X	
		24	BS	N		X
		12	BS	N		X
		2	BA	N		X
		38	MA	Y		X
		32	BS	N		X
		18	BA	N		X
		3	BS	N	X	
		18	BS	N		X
		30	MA	Y	X	
		12	BA	N		X
		17	BS	N		X
		30	MA	Y		X
		9	BS	N		X
		28	MA	Y	X	
		11	BA	N		X
		6	MA	Y		X
		16	BA	N		X
		6	BS	N	X	
		29	BS	N	X	
		22	BA	N		X
		6	BA	N		X
		23	BA	N		X

Appendix Z

District Permission Letter

[redacted]

Superintendent

[redacted]

To: Northwest Nazarene University – HRCC Department.

[redacted] which is part of [redacted] is allowing Greg Wiles (Doctorate Student from NNU) to complete his research study using our teachers during the fall semester of 2012. We are giving Greg Wiles permission to use the data that is collected from the surveys administered in his dissertation. We are also allowing Greg to conduct a staff development session with the [redacted] staff.

[redacted]

Principal

[redacted]

Superintendent

[redacted]

Appendix AA

E-mail Artifact

On Thu, Nov 1, 2012 at 6:47 PM, Wiles, Greg <gwiles@nsd131.org> wrote:

Are those lessons ok?

Greg Wiles

Principal

Lone Star Middle School

From [redacted]m
Sent: Thursday, November 01, 2012 9:29 PM
To: Wiles, Greg
Subject: Re: questions

Lesson 1 went great. Staff is ready to make serious changes. Hopefully we will get district support in making Power School work for us.
--
[redacted], Principal
[redacted] Middle School
cell [redacted]
office ([redacted]

Appendix BB

Behavior and Engagement Rubic

Behavior Rubric

	4 consistently	3 Frequently	2 Occasionally	1 Rarely
Respect				
▪ Follows teachers instructions				
▪ Talks quietly and appropriately				
▪ Keeps hands to self				
▪ Respects others				
▪ Uses appropriate language				
Relationships				
▪ Maintains Healthy friendships				
▪ Makes responsible choices				
▪ Resolves conflicts effectively				

Engagement Rubric

	4 consistently	3 Frequently	2 Occasionally	1 Rarely
Responsibility				
▪ Completes homework & class work on time				
▪ Stays on task and manages time				
▪ Participates in class activities; offers ideas, asks questions				
▪ Persists in problem solving				
Readiness				
▪ Brings materials to class				
▪ Arrives to class on time				

Appendix CC

Power of I

Summary:

The Power that comes from the Power of I can be gained or lost by how teachers implement the program. There are nine central components of the Power of I. Those with an * are nonnegotiable and must be implemented for the program to be successful. Others are components that schools have used that have worked for them, but not for other schools. Each teacher or each school will have to tweak the components of this grading practice in order to make it work for them. If this practice is rolled out throughout the whole school, it is important that it look the same from class to class.

1a. Students no longer receive zeros when work isn't turned in; they don't have an option not to turn in work.

- **Teachers have made this clear from the beginning of the year.**
- **Teachers have other "consequences" for work not done, not done completely, or not done satisfactorily.**

1b. Teachers no longer assign grades below a C; students are required to redo/revise work to get it at least to the 'basic' or 'C' level.

2. Late work is just that—late—but it must be *completed* if teachers are to correctly determine if students know and understand the standards being taught and assessed.

3a. Students <u>must be given extra help opportunities</u> *(required)* to complete the work during the school day.

- **before or after school (never during your class, ever),**
- **Saturday School,**
- **or whatever fits your school's possibilities.**

(This piece is completely up to schools to determine how this help can best be delivered.)

3b. Some schools require students to attend extra help prior to turning in any late or redo work. (contact Valerie Carrier, principal, Rivera Middle School, Dade County, Florida)

4. Consequences change for students not having work ready to turn in on time:

- **Must contact students' parents** and solicit their assistance—this must begin early and will have the greatest impact.
- **Requires a parent conference** at a predetermined number of missed assignments or failed tests, etc.

Appendix CC (cont.)

- **Require students to attend extra help after school** in an extra help setting to complete work (this takes some work, but many schools are having great success with this).
  - **Some schools require all students to attend extra help** prior to turning in any late assignment or redo work.
  - **Some policies state that students are not allowed to participate in any extracurricular activity** (sports, band, chorus, clubs, events, and practices) if they are missing any assignments or have attempted to turn in poor-quality work.

5. **Students will receive an I** on any assignment not turned in, and ***<u>One</u>* or more I's will result in an I grade for any report period.**
   - **Teachers must become more judicious about the kinds of assignments they give and assign a grade to.**
   - **Which assignments are worth "chasing down"**

6. **Tests may be excluded from the policy.**
   - **Teachers may *<u>require</u>* all students who do not pass a major test to *<u>retake</u>* the test during extra help time until they make a passing grade.**
     - **If so, the highest grade to be recorded in the grade book should be the lowest C grade.**
       - **Caution: students may choose to not study and simply retake tests if they think they can get a higher A or B grade.**
   - **Teachers *<u>may</u>* allow any student to retake a test for a higher score, but this is not a part of Power of I.**
   - **Teachers *<u>may</u>* choose to give students *opportunities* to raise test score to a C by coming in during extra help times to retake tests.**

7. **Students cannot receive an A** (or a B in some schools) on any assignment that is late or turned in incomplete (some schools have instituted specific time periods)

8. **Students never receive an F *<u>if</u>*** an assignment is completed within the year or semester (determined by each school as appropriate).

9. **A few students will still fail** no matter what you do**. So…**
   Final report cards have asterisk or note reporting to parents that the F is a result of failure to complete work.

The goal is to get all groups of students to meet course standards at an acceptable level. Knowing that it WILL NOT WORK for all students, each school should set goals by tracking the success of this program using appropriate data. First,

- **% of students currently making zeroes =**

Then,

- **% acceptable after first semester =**
- **% acceptable after first year =**