2011

Grading Practices: Teachers' Considerations of Academic and Non-Academic Factors

Diana Yesbeck
Virginia Commonwealth University

Follow this and additional works at: https://scholarscompass.vcu.edu/etd

Part of the Educational Leadership Commons

© The Author

Downloaded from
https://scholarscompass.vcu.edu/etd/2586

This Dissertation is brought to you for free and open access by the Graduate School at VCU Scholars Compass. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of VCU Scholars Compass. For more information, please contact libcompass@vcu.edu.
Grading Practices: Teachers’ Considerations of Academic and Non-Academic Factors

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University.

by

Diana Marie Yesbeck
Bachelor of Science, Virginia Commonwealth University, 1988
Master of Education, University of Virginia, 2004

Director: James McMillan, Ph.D.
Professor, Foundations of Education
School of Education

Virginia Commonwealth University
Richmond, Virginia
December 2011
Acknowledgements

My greatest support system is my family. They followed me through this journey, and each member provided love, encouragement, and guidance to ensure that I would persevere and complete this quest. To finish this project, the sacrifice of time with my son Jake and my extended family was the most difficult to endure, yet the inspiration they afforded me fills my heart with pride and love for their endless support. I wish to express my sincerest appreciation to Jake, my parents, Vera and Edward, my siblings, Donna, Greg, Ed, and Nancy, and their children, Katherine, Luke, Gregory, and Adam. This would not have been possible without their love and encouragement.

Of equal importance is my sincere gratitude to my dissertation chairperson, Dr. James McMillan, and my committee members, Dr. Cheri Magill, Dr. Paul Gerber, and Dr. William Bosher. Along with Dr. Leslie Bozeman, my committee generously shared their experiences, their wisdom, and their guidance.

Finally, friends, who are expert writers, assisted the process when they offered to review the work in stages, providing editing suggestions. Their ideas improved the overall product, and to them, I am grateful.
Table of Contents

List of Tables .................................................................................................................. vi

Abstract ........................................................................................................................ v

Chapter I: Introduction .................................................................................................... 1
  Statement of the Problem ......................................................................................... 3
  Purpose of the Study ............................................................................................... 4
  Rationale and Significance of the Study ................................................................. 5
  Literature/Research Background ......................................................................... 6
  Methodology ........................................................................................................... 9
  Summary .................................................................................................................. 10
  Definition of Terms ............................................................................................... 11

Chapter II: Review of Literature .................................................................................. 13
  Measurement .......................................................................................................... 14
  Assessment ............................................................................................................. 18
    Formative and Summative ................................................................................. 20
  Grading .................................................................................................................. 22
    Purposes of Grading ........................................................................................... 24
    Grading Practices................................................................................................ 28
    Influence of Electronic Grading Systems .......................................................... 34
    Standards-Based Grading .................................................................................. 37
    Grading Policies ................................................................................................. 38
  Good Practice and Poor Practice in Measurement, Assessment, and Grading to
    Reflect Student Achievement and Accountability ........................................... 40
  Summary .................................................................................................................. 43

Chapter III: Methodology ............................................................................................. 45
  Research Questions ................................................................................................. 46
  Design ..................................................................................................................... 46
  Rationale for Design ............................................................................................... 47
  Participant Selection and Setting ......................................................................... 48
  Pilot Test ................................................................................................................ 50
  Instrument Revision ............................................................................................... 50
  Data Collection ....................................................................................................... 51
  Data Analysis .......................................................................................................... 52
  Delimitations .......................................................................................................... 53
Recommendations for Research ..............................................................132
List of References ..................................................................................135
Appendices ............................................................................................143
  Appendix A: Interview Guide ...............................................................144
  Appendix B: Measurement Theory Recommendations .........................146
  Appendix C: Research Subject Information and Consent Form ..............147
  Appendix D: Letter to Principals ..........................................................151
  Appendix E: Letter to Teachers ............................................................153
Vita .............................................................................................................155
List of Tables

Table 1. A Comparison of the Purposes of Grading from the Literature to Date ........ 30
Table 2. Comparisons of Good Practice and Poor Practice in Grading ..................... 42
Table 3. Teacher Demographics ............................................................................ 57
Table 4. Code Frequency ...................................................................................... 62
Table 5. Data Analysis Themes and Subthemes ...................................................... 63
In the current era of student accountability, coupled with high-stakes testing, schools have focused on the alignment of standardized curriculums and assessments. However, developing standardized grading practices are still under examination. Grading serves as an important responsibility, yet many teachers still find the process a challenge of determining which academic and non-academic factors correctly represent student achievement.

This qualitative study was designed to examine the grading factors teachers consider when determining student final grades. Middle school language arts teachers from one mid-sized suburban school district in Virginia participated in qualitative interviews. The interviews addressed the following topics: (1) the purposes of grades, (2) the grading factors teachers consider when they grade, (3) the teachers’ influences in
determining their practices, and (4) how teachers’ grading practices relate to measurement theory.

Overall results in the areas of the purpose of grades, the use of academic and non-academic factors, teachers’ influences, the use of formative and summative assessments, and the need for professional development are consistent with the literature. With respect to how teachers’ practices relate to measurement theory, the findings are consistent with the literature from previous studies. This suggests that although measurement theory experts recommend that achievement factors should be the only factors that determine student grades, the results of this study indicate that teachers use a mixture of variables in determining student grades, known as hodgepodge grading.

Recommendations for practice include teacher reflection on determining why they grade, becoming familiar with measurement theory recommendations in terms of using academic factors that measure student achievement, how to use non-academic factors in other ways to support student learning, and providing staff development in the area of grading practices and how practices support measurement standards.
Chapter I

INTRODUCTION

With an increase in student and teacher accountability, coupled with mandated state assessments to meet the requirements of the *No Child Left Behind Act of 2001*, a significant amount of literature exists which focuses on the variation of grading practices. In addition, the *Race to the Top* (RTTT) federal program, under the American Recovery and Reinvestment Act of 2009, is intended to reward states for educational innovation by creating standards and assessments that prepare students for success in college and in the workplace. In order to meet the requirements, standardization of curriculums and assessments are a focal point, yet developing standard grading practices are still under examination. Although measurement experts recommend a variety of grading practices, teachers develop their own practices based on the importance they place on academic and non-academic factors. Since grades communicate results to students and parents, different messages are being conveyed about academic performance and indicators of student achievement (McMillan, 2001).

Guskey (2004) believes that grading is an important professional responsibility, yet serves as a great challenge to teachers since few have had formal training in grading methods. In addition, many teachers have limited knowledge about the effectiveness of various grading practices. In Guskey’s 2006 study of understanding why and how teachers grade, the results showed that educators draw primarily from their own
experiences as students in determining the grading practices they employ. Guskey maintains that educators must make every effort to ensure that grading practices are clear, explicit, and objective as possible. Furthermore, he believes educators need to refrain from using personal opinions and unconscious biases as influences while assigning grades, ensuring that above all, their grading policies and practices will be equitable and unprejudiced.

Grading practices continue to be controversial and misunderstood. A disconnect exists between grading purposes, practices, and policies and the current era of accountability. Student performance in class should, in theory, partially match performance on standardized tests. Measurement experts believe that since educational reform is on a rise in the areas of curriculum development, standardized assessment, and high-stakes testing, there is a greater need to further the reform to include grading practices among classroom teachers, in order to eliminate the attitude that since grading practices are so ambiguous, that the actual grade is meaningless (Marzano, 2000).

Teachers’ grading practices are becoming more scrutinized to adapt to more accountability testing and performance-based assessments and determining how to grade the assessments that accurately measure student achievement. Given the variety of grading practices that studies and literature report, there is a need to more fully understand why variation exists among teachers’ grading practices and how the practices relate to measurement theory using qualitative interviews. Inquiries may lead to understanding a summary of grading practices that teachers use to generate grades, such as the measurement procedures they use, their rules of evidence or the standards they apply (Stiggins, Frisbie & Griswold, 1989).
This dissertation reports the results of a qualitative study which investigated the academic and non-academic factors that teachers consider when determining grades. Academic factors simply include student performance. Non-academic factors include responsibility, effort, attitude, behavior, motivation, and attendance. The study used in-depth interviews with ten teachers to determine the reasons for assigning or calculating final grades, and was limited to middle school language arts teachers.

**Statement of the Problem**

The problem this study examined is complex in nature because teachers’ judgments and values determine current grading practices. According to Brookhart (1994), grades are the most common form of interpreting measurement results, yet the practice of grading varies greatly. In 2006, Guskey affirms his idea that grades have been identified as example of unreliable measurements. Since teachers’ criteria differ, significant variation remains in the grading practices of individual teachers.

Brookhart (1993) believes that students’ grades should reflect the relationship of a student’s ability and their expected performance, contrary to Wormeli (2006) and Winger (2005) who believe that grades should reflect an accurate measure of student mastery of content. Additionally, Winger and Wormeli believe that educators ultimately fail students when non-academic factors, such as effort, responsibility, and attitude are misused when calculating grades because students are learning that diligent work and cooperation are the components of their grade, not the ability to make connections with the learning through deep understanding. According to McMillan (2001), there is much variability in grading practices, which offers little consistency across schools and within classrooms, even when schools and teachers adopt the same grading guidelines. The
result of McMillan’s study showed that teachers employ a variety of grading practices that can be summarized into four distinct components: (1) academic achievement, (2) academic enablers, such as effort, ability, improvement, and participation, (3) use of external benchmarks, and (4) use of extra credit and borderline cases. Consistent with this study were findings from previous studies (Frary, Cross & Weber, 1993; Stiggins and Conklin, 1992; Stiggins, Frisbie & Griswold, 1989) that academic achievement continues to be the most important component in grading students. However, consistent with Cross and Frary (1999) are the importance that teachers place on non-academic factors such as effort, participation, and improvement.

**Purpose of the Study**

The purpose of the study was to examine the academic and non-academic factors teachers consider in their grading practices to gain a better understanding of the discrepancy within teachers’ grading practices. Academic factors include student performance and achievement, while non-academic factors include a variety of items, including, but not limited to: responsibility, effort, attitude, behavior, motivation, attendance, chemistry between teacher and student, class size, and teachers’ efficacy in working with different student populations, such as those in special education. The results of this study will assist with the understanding of middle school language arts teachers’ approaches to assigning grades in an effort to understand the arbitrary grading practices teachers employ. It informed practice by bringing awareness to teachers’ understanding of measurement experts’ recommendations, and it produced meaningful conclusions based on similarities and differences on how teachers determine student grades. Furthermore, the results of this study will help enable additional research for
other content areas to determine variation among grading practices and why the variation exists.

The research questions served as a guide for the researcher. The questions focused on which academic and non-academic factors teachers consider and why these factors were chosen. The four major questions included:

1. What influences teachers’ grading practices?
2. What academic factors were considered in determining the grade, and what beliefs and/or values were used to determine those factors?
3. What non-academic factors were considered in determining the grade, and what beliefs and/or values were used to determine those factors?
4. What gaps exist between teachers’ grading practices and recommendations made by measurement experts?

Rationale and Significance of the Study

This study has contributed to the literature by developing an understanding of the various ways grades can have different meanings. The current literature indicates that the study of grading practices will add to the understanding of the significance and meaning of student grades. Literature in this field continues to provide a variety of information while simultaneously revealing unanswered questions and the need for further research (McMillan, 2001). Determining the academic and non-academic factors used in teachers’ grading practices will afford teachers and educational leaders with unique challenges that have not been fully explored nor examined. The conclusions and significance of this study may lead to implications in constructing staff development to determine more fair and best practices in recording student achievement in terms of the relationship with
academic and non-academic factors with an emphasis on student accountability (Allen, 2005). Allen supports the idea of requiring teachers to participate in teacher-education programs that emphasize adequate instruction on classroom assessment and grading practices, and further supports that measurement textbooks need to address grading practices in which teachers can fully understand the importance of improving practices that relate to the purposes of grades.

**Literature/Research Background**

The terms *measurement, assessment, and grades* are often linked and used synonymously to describe the evaluation and review of student work. Although they are closely linked, there are specific distinctions among the three terms.

McMillan (2008) describes *measurement* as a process used to quantify the degree of how much something has been demonstrated with a value that is assigned, based on a scale. Teachers measure traits, such as performance, behaviors, and attributes, by administering tests and recording the measurement of the trait. Measurement experts believe that grades should be focused on levels of student achievement, but recognize that other factors are considered when assigning grades. The other factors include effort, progress, participation, behavior, and attitude. When other factors are considered, measurement experts question the validity and reliability of the grades (Stiggins, 1999). This is supported by Guskey (2006) who believes that the use of arbitrary grading practices questions the validity of a student’s grade. Gallagher (1998) believes that the most fundamental principles related to measurement and classroom grading are validity and reliability. Furthermore, measurement experts’ recommendations (from a sample of introductory measurement textbooks), as outlined in Stiggins, Frisbie & Griswold’s
(1989) study, believe that most current grading practices are not in line with best practices for student assessment. The most conclusive recommendation that is not consistently followed among teachers is the use of assigning grades based primarily on student achievement. Other non-academic factors, such as effort, participation, interest, and attitude, are often used by teachers to determine a final outcome.

Wormeli (2006) is a strong proponent of assigning grades based only on academic factors. He believes that teachers mistakenly use some grading practices as accountability measures. The three practices that Wormeli questions are (1) how teachers assess homework, (2) how teachers use zeroes as a measure of accountability, and (3) teachers’ willingness to allow students to retake tests to demonstrate mastery. If students were not responsible in their academics, all three of these practices would be detrimental to students’ grades; however, Wormeli contends that incorporating non-academic factors into academic indicators of mastery is not helpful or accurate, nor does it teach students to be more responsible and thereby accountable. Therefore, he proposes other ways to use assessment and feedback to teach accountability which does not conversely relate to poor academic grades in terms of student achievement.

Airasian (1997) describes assessment as a way to help with decision-making by collecting, synthesizing, and interpreting information. It differs from former views of assessment which narrowed the definition to the mere task of testing and gathering information. Grading is based on professional judgment, and McMillan (2008) finds the practice to vary significantly. It is the method by which grades are determined that necessitates a deeper understanding of what researchers define as the purpose of grading.
In order for grading practices to improve and be more accurate and educationally meaningful, Allen (2005) suggests that two major changes occur. The first change is to determine student grades based solely on achievement, without factoring in non-academic factors, such as effort, attitude, responsibility or behavior. Secondly, Allen suggests that teachers need professional development in terms of creating effective assessments to accurately measure student achievement objectively. For example, teacher preparation courses in measurement theory, educational psychology, and methodology could effectively provide training in teacher-education programs in the areas of classroom assessment and grading practices.

In an effort to determine why teachers have difficulty assigning grades objectively, Tombari and Borich (1999) narrow the reasons to three. The first discrepancy teachers have with varying grading practices is tied to academic achievement and non-academic factors, such as responsibility and attitude. Second, teachers have difficulty reporting one single grade (as in a final report card grade) for students when the grade represents multiple components (which may be an invalid measure of the student’s performance) supporting Guskey’s (2004) findings of how teachers defined grades. Finally, teachers are, for the most part, inadequately trained in the areas of assessment and grading, a supposition which further supports Allen’s (2005) theory about effective professional development.

O’Connor (2007) believes that communication at the school and district levels needs to take place to develop a shared vision concerning the primary purpose of grades and defining student achievement as a measurement of performance. Often, grading and assessment items are overlooked for annual review because of the perception that there is
rarely change. However, O’Connor supports the idea of an annual review of school division’s handbooks to ensure that all stakeholders understand student achievement as defined by measurement practices, known as grading. Grading is a major responsibility of classroom teachers, yet many teachers find the practice uncomfortable and difficult due to the number of factors that may be considered in assigning grades (Allen, 2005).

Methodology

To ensure a meaningful study, qualitative in-depth interviews of ten middle school language arts teachers were used. A phenomenological study was appropriate to focus on one shared phenomenon for an in-depth understanding. A purposeful sampling method was used in this study, which is the most common type in qualitative educational research (McMillan and Schumacher, 2006).

Participants were limited to teachers who teach middle school language arts from one suburban central Virginia school district. Language arts teachers, who use much subjectivity in their grading, were selected to complete the qualitative interviews to keep the comparison of teachers’ grading practices within one core subject. Data collection involved asking teachers to participate in qualitative in-depth interviews in which the interview protocol focused on what influences teachers’ grading practices, what academic and non-academic factors teachers used in grading, and how their grading practices related to measurement experts’ recommendations.

The participants of the qualitative study were required to give consent to participate in the initial study. Once the interviews were transcribed, a general analysis took place in the discovery phase. Using codes, the deconstruction phase began based on categorizing recurring ideas and major themes. Results from the qualitative data analysis
are shared in narrative form with emphasis placed on the comparison and contrast of the participants’ responses in Chapter four (McMillan & Schumacher, 2006).

**Summary**

Research supports that the task of assigning grades vary, and that there are numerous grading practices (Brookhart, 2009). Although school systems may employ standards in terms of grading, teachers usually use arbitrary methods and a combination of both academic and non-academic factors to determine grades. Since there does not appear to be a definitive method in calculating grades, grading practices continue to be controversial and misunderstood. Wormeli (2006) is a strict proponent of using only academic factors in calculating grades. Conversely, McMillan (2008) recognizes that teachers’ subjective evaluations and intuition are valid contributing factors used in grading.
**Definition of Terms**

Within the context of this study, the following definitions will be used:

*Academic Factors* – those that are considered in grading practices – student achievement or performance which demonstrates mastery of content (Wormeli, 2006)

*Achievement* – mastery of knowledge and skills (Griswold, 1993)

*Aptitude* – the capability of a student and how he/she puts his/her capability to use (Griswold, 1993)

*Assessment* – a way to help with decision-making by collecting, synthesizing, and interpreting information (Airasian, 1997)

*Attitude* – degree to which students communicate interest and willingness to learn (Griswold, 1993)

*Effort* – degree to which students work to get the task done (Griswold, 1993)

*Formative Assessment* – assessment method that provides ongoing feedback and helps students make informed decisions (Popham, 2008; Brookhart, 2009; McMillan, 2008)

*Grades* – a score or mark based on a quantity of measurement (McMillan, 2008)

*Grading Practices* – the methods teachers use to assign student grades (Brookhart, 2009)

*Measurement* – a process used to quantify a degree of how often something has been demonstrated and then a value is assigned (McMillan, 2008)

*Non-Academic Factors* – those that are considered in grading practices – factors that relate to student behaviors, work habits, and attitudes (Brookhart, 2009)

*Student Accountability* – the degree to which students are responsible for their learning (McMillan, 2008)
*Summative Assessment* – assessment method that provides a final grade as an outcome of learning (Brookhart, 2009)
A review of the background knowledge and literature in terms of grading practices and student performance is the focus of this chapter. The information contained is organized into headings of measurement, assessment, and grading. Within the section devoted to grading, additional information includes the purpose of grading, grading practices, grading policies, electronic grading systems, and standards-based grading.

In order to gain information about the topics, a search for literature took place using the VCU Library website, searching in the databases of ERIC, Academic Search Complete, Education Search Complete, PsychArticles, PsychInfo, and the Psychology and Behavioral Sciences Collection. Only peer-reviewed articles were part of the search. In addition, criteria including article dates since the year 2000 were used to search for the articles.

Keywords included grading practices, assessment, measurement, in addition to research studies, and performance. The research studies were limited to articles that involved K-12 teachers in public schools in the United States. The standards by which the search of literature was vetted included using initial sources from the review of primary sources and empirical studies, which often led to other sources, such as books,
textbooks and secondary sources related to grading practices, listed from the original primary sources of research.

**Measurement**

Often the words *measurement, assessment and grades* are used interchangeably. In some way, all three terms describe evaluation. McMillan (2008) describes measurement as a process used to quantify the degree of how much something has been demonstrated and then a value is assigned which is based on a scale. Teachers measure traits, such as performance, behaviors, and attributes, by administering tests and recording the measurement of the trait.

In terms of measurement, a grade should reflect a clear measure of the best a student can do. However, grades reflect a mixture of multiple factors, resulting in an ineffective communication system (Tomlinson and McTighe, 2006). Brookhart (1994) agrees that since grades are the most common form of reporting student performance, developing consistent grading practices are imperative for interpreting the meaning of grades. To better understand the meaning of student grades, O’Connor (2007) believes that educators should determine that grades assigned to students are consistent and accurate as a clear measure of the students’ understanding, and the grades reflect the school’s content standards and desired learning outcomes.

Research supports that a gap clearly exists between current grading practices and measurement theory, which is documented in Brookhart’s comparative study of teachers’ grading practices. In Brookhart’s report, grading practices from 19 different studies were compared to recommendations from measurement experts. It was determined that although measurement experts have outlined recommendations for grading
(recommendations are listed on the next page), teachers employed a mixture of grading practices, which met some of the recommendations, but not all. The greatest discrepancy fell in the area of measuring student achievement in terms of academics as well as non-academic factors, such as effort, progress, participation, behavior, and attitude. However, measurement experts caution against the use of non-academic factors and arbitrary grading practices because engaging in this practice questions the validity and reliability of the students’ overall grades (Gallagher, 1998; Guskey, 2006; Stiggins, 1999).

Gallagher’s (1998) recommendations are based on basic measurement theory, assessment reliability and validity, and designing assessments that bridge the gap between research and practice in the area of student assessment. Guskey’s (2006) study found that establishing grading policies and practices need to focus on the importance of clarity and fairness that are not perceived as ridiculing or embarrassing to students. Stiggins (1999) found that classroom assessment training in teacher education programs lacked clarification of grading practices and policies, questioning the validity and reliability of classroom assessments.

In one of the studies that was reviewed, Stiggins, Frisbie and Griswold’s (1989) case study surveying 15 secondary teachers, 19 dimensions of grading practices were used as a framework. Of the 19 recommendations from measurement texts, the findings concluded that teachers were consistent with eight of the recommendations and were in discrepancy with 11 of the recommendations. Below are the eight recommendations made by measurement experts in terms of grading practices that were consistent with teachers’ practice.

1. Grading methods should be communicated to students.
2. Student attitudes should not be used for grading purposes.

3. Students’ levels of interest in the subject matter should not be used as a grading component.

4. Students’ personalities should not be used in assigning grades.

5. Written tests should be used as a primary means of measuring achievement.

6. The use of oral questioning is effective in monitoring learning, but should not be considered in grading.

7. Performance assessments are appropriate to measure student performance and include in grading methods.

8. Most class sizes are too small to expect that achievement will be distributed normally; therefore, it is recommended that the normal distribution not be used as the standard for judging the appropriateness of the distribution of grades.

   The remaining 11 recommendations were met with discrepancy among teachers’ grading practices as found in the same study.

9. Achievement should be the sole ingredient in determining grades.

10. Intelligence and cognitive ability should not be used in determining grades.

11. Although effort may impact achievement, it is recommended that effort not be a component of grading. Feedback should be reported separately.

12. Formative-type of assessments should not be used in grading.

13. Amount of grading data gathered should reflect enough evidence over time for each student to master the material. It is recommended that two or three high-quality assessments per quarter be used in determining grades; however,
smaller formative assessments should not be included in determining the final grade.

14. Teachers must evaluate the quality of data gathered, in terms of reliability and validity.

15. Consistent grading policies need to be written and distributed within school districts.

16. Aggregating grading components are necessary to create the composite index of achievement.

17. Fixed percentages are recommended for setting grade cutoff scores.

18. Using total point accumulations is not recommended because the arbitrary cutoffs are set that has no clear reference and gives ambiguity to the meaning of the grade.

19. In borderline cases, only student achievement data should be used, not non-achievement factors.

Stiggins, Frisbie and Griswold (1989) determined three general possibilities for the discrepancies in what measurement experts agree as best practices in grading as opposed to what teachers employ in the classroom. First, the term best practice may mean different things to different teachers. Best practice may be based on opinion or philosophical positions rather than factual knowledge. Secondly, measurement experts are far removed from the classroom and fail to take into account the constraints of the practitioner in the classroom. Thirdly, teachers may be unaware of the recommendations or have had little training in the area of measurement theory. Stiggins, Frisbie and
Griswold conclude by encouraging more in-depth research in the area of aligning grading practices with measurement theory.

McMillan’s 2001 study with secondary teachers, grades 6 through 12, indicated that teachers have implemented some of the expectations for assessment reform based on measurement experts, while maintaining some traditional practices in others. In another study, Stiggins and Bridgeford (1985) explored teacher-developed assessments, and the study yielded two major findings: (a) measurement training has left teachers poorly equipped to design their own form of assessment and (b) measurement training did not allow for teachers to make daily informal observational assessments to address educational decisions that teachers need to make on a regular basis.

Conversely, in Brookhart’s (1994) comparative summary, three possible reasons were cited for why teachers did not follow recommended measurement practices. First, teachers may adopt the attitude that recommended practices are opinion-based and not fact-based. Second, measurement recommendations do not realistically fit the classroom setting. Third, teachers lack measurement training. Although measurement, assessment, and grades are closely linked, there are distinctions among the three terms. A closer look at assessments will assist in understanding the tools teachers use to measure student learning.

Assessment

Classroom assessment can fundamentally transform classroom instruction (Popham, 2008). Former views of assessment constrict the task of assessment to testing and gathering information (McMillan, 2008); however, current views describe
assessment as a way to help with decision making by collecting, synthesizing, and interpreting information (Airasian, 1997).

Popham (2008) supports the use of assessment by evaluating student work frequently to make instructional decisions. In fact, instruction and assessment are inseparable, and can be thought of as a cyclical process. McMillan (2008) describes the transformation that assessment has undergone in recent years, from determining student skills as isolated and disconnected facts to interconnected skills that teach students how to integrate knowledge. McMillan’s description continues to outline the transformation from a student’s basic and surface-level understanding to deeper knowledge and application. Furthermore, McMillan describes current assessment practices as continual and motivational, rather than sporadic and controlling. This is contrary to some practices, due to the development of current national and state assessments which lack short answer questions, essays, and performance-based tasks.

Due to these types of cursory assessments, coupled with the era of high-stakes testing, Popham (2008), Brookhart (2009) and McMillan (2008) strongly recommend that teachers employ more informal types of assessment on a regular basis to balance the types of assessments in the classroom. Popham supports many informal and formative style assessments that provide meaningful feedback for the students and the teacher in order to make informed instructional decisions. As a proponent of formative assessment, Brookhart agrees that formative assessments assist with the development of knowledge and skills. McMillan recommends in providing ongoing feedback through informal assessments, such as observations, journaling, and individual conferences assist with the overall understanding of students’ understanding of incremental learning steps.
Formative and Summative

To better understand the purposes of assessment, it is helpful to understand the two types of assessment that research currently supports – formative and summative. As early as 1969, Bloom used the term *formative evaluation* in identifying a stage in the teaching-learning process. By comparing the formative years of individuals – those years that shape people – to formative evaluation, Bloom believed that formative evaluations shape whatever is being evaluated. The purpose of formative assessment, as defined by Popham (2008), Brookhart (2009) and McMillan (2008) is to gather evidence to help students make informed decisions about how to improve their current learning. Taras (2005) clarifies the definition further to include that formative assessment indicates the existence of a gap between the actual level of the work being assessed measured against the standard. Simply put, formative assessment provides meaningful feedback to students to assist in improving their performance.

Sadler (1989) created a theory that formative assessments help students identify the gaps between the student’s current level of understanding compared to the criteria that meets satisfactory or excellent performance. The feedback provided in formative assessments supplies specific information to close the gap of understanding, performance, and achievement. In order to close the gap, Sadler suggests three key points for students: (a) have a clear understanding of the concept and what quality performance looks like in mastering the concept; (b) spend valuable time in reflection in comparison of the one’s performance to the concept while it is happening; and (c) employ a variety of learning strategies to close the learning gap. Additionally, Sadler recognizes that teachers play a
critical role in helping students by providing effective and meaningful feedback about their progress and performance.

In support of Sadler’s (1989) recommendations, Ames and Archer (1988) also suggest that teachers can foster a classroom environment of continuous improvement by emphasizing six strategies: (a) help students understand that success in the classroom is measured by improvement and progress, not a final grade; (b) place importance on learning and hard work; (c) focus on how students are learning; (d) recognize that errors and mistakes are part of learning; (e) place importance on the process of learning; and (f) evaluate students’ work in progress.

McMillan (2008) describes formative assessment as a cyclical teaching tool. During instruction, formative assessment allows students to ask questions, develop intrinsic motivation to deeply understand the material, and apply the new learning. Formative assessments assist teachers in making sound decisions concerning instruction. For example, when students perform poorly on a homework assignment or quiz, teachers can monitor and adjust the instruction according to the individual’s weaknesses.

If formative assessment is intended as feedback to improve, conversely, the purpose of summative assessment is to provide a final grade as an outcome of learning (Brookhart, 2009). Summative assessments include unit tests, standardized tests, or any assignment that counts towards the report card grade. Unfortunately, summative evaluations can produce a final outcome that does not allow student and teacher dialogue about ongoing instruction; however, Wolf (1993) believes that good teachers can turn summative evaluations into teachable moments. For example, item analyses in summative assessments produce meaningful data for teachers. When the same question
appears to be problematic for a large percentage of the students, the teacher can determine if the instruction matched the learning outcome, making it a teachable moment for the educator. If, in fact, the question measured an appropriate learning outcome and the students performed poorly, it gives the teacher another opportunity to adjust instruction accordingly, making it a teachable moment for the students. In addition, the teacher also has the opportunity to check the validity of the question in terms of matching curriculum standards. In Fleming and Chambers’ (1983) study, teacher-made tests were analyzed. Used as summative evaluations, these tests were limited to lower levels of cognitive thinking, relying on students’ recall, knowledge, and comprehension. Components of analysis and synthesis were rare in the summative evaluations.

Summative assessments, such as standardized tests, measure students’ understanding of material over a period of time and the data are often used to help schools make decisions about instructional programs. Additionally, summative assessments are used in the reporting of grades on individual student reports (Brookhart, 2009: McMillan, 2008). Brookhart suggests caution in the use of making educational decisions based solely on summative evaluations, because students would be expected to move forward instructionally, whether or not they mastered or demonstrated previous material, which contradicts the art of meaningful learning.

Grading

Grades mean the actual score or mark that a student receives based on the quantity of measurement and decision-making of assessment. McMillan (2008) describes grades as a means of communication concerning student performance. In recent studies, results found that the meaning of grades continues to be controversial. In fact, Cross and
Frary (1999), explain that school marks and grading have been a source of controversy since the turn of the century. Furthermore, Marzano (2000) believes that there has been widespread speculation on the value of grades due to an inaccurate system of grading practices. Therefore, when the three components – measurement, assessment and grading – are aligned, then the grade reflects a true indication of the student’s understanding on an assessment that was developed based on measurement recommendations.

There are three types of grading, according to Brookhart (2009). Criterion-referenced grading is comparing student’s work to standards – these are the type of state assessments that meet the requirements of No Child Left Behind. Second, when grading on the basis of comparing student’s work to their previous work to show progress, this is known as self-referenced grading. The third type of grading, known as norm-referenced grading, is when a student’s work is compared to the work of others.

The validity of grades is called into question when the grades do not represent a true measure of the learning outcome. To ensure validity, it is recommended that performance on assessments match instructional goals and objectives (Brookhart, 2009). McMillan (2008) believes that grading is based on professional judgment, and finds that the practice varies considerably among teachers, even teachers in the same school. For example, if the three components of measurement, assessment, and grading are not aligned, the student’s grade may or may not reflect student understanding on a summative assessment that did not follow measurement recommendations – thereby making the grade meaningless and arbitrary.

Griswold’s 1993 study included scenarios based on student performance, ability, effort and attitude. The teachers rated two scenarios, determining if the students should
receive a higher or lower grade based on the four categories. From this study, Griswold concluded that grading involves one or more of four factors: (a) achievement – mastery of knowledge and skills; (b) effort – the degree to which students work hard to get the job done; (c) attitude – the degree to which students communicate interest and willingness to learn; and (d) aptitude – the capability of students and how they put it to use.

The history of grades in American education is not precise, but one aspect that most historians agree on is that prior to the late 1700s, students were not given grades but were given meaningful feedback on their performance in the form of narrative comments (Marzano, 2000). In current American education, Brookhart (2009) believes that while grades about student achievement should be based on solid, high-quality evidence, descriptive feedback adds valuable information concerning the student’s level of performance and achievement.

**Purposes of Grading**

Brookhart (2009) believes that in an ideal world, grades would be replaced with discussions of performance, emphasizing strengths and weaknesses, in order to continue to improve until a task or skill is mastered. However, in American education, grades have become the standard by which many students, schools, and programs are compared. Grades are essential in communicating to students and parents about their achievement of learning goals, yet grading policies and practices lack consistency. Brookhart believes that for grades to be meaningful and systematic, a method in which grades are determined necessitates a deeper understanding of what researchers define as the purpose of grading. As described by Brookhart, the purposes of grades are threefold and date back for at least 50 years:
1. To rank or sort students, possibly for higher education recommendations.
2. To report results as in communicating to parents and students.
3. To contribute to learning by providing feedback and motivating students.

However, in an earlier study, Brookhart (1993) stated that students’ grades should reflect the relationship of a student’s ability and their expected performance, indicating that student’s cognitive ability may match student achievement. This is contrary to Wormeli (2006) and Winger (2005), along with many other researchers, who believe that grades should only reflect an accurate measure of a student’s mastery of content. In support of Wormeli and Winger’s beliefs, Sadler (2010) agrees that grades should reflect an authentic representation of a student’s level of academic achievement, and if a grade is to be trusted to communicate this achievement, then the only factors that should contribute to the grade must qualify as achievement. In addition, Sadler believes that determining whether a particular component qualifies as achievement is a classification issue rather than a measurement issue.

According to Guskey (2004), when teachers were asked to define the purpose of grading, they indicated that both academic and non-academic factors needed to be considered when calculating grades. The teachers Guskey identified in his research determined that the purposes of grading fell into six categories: (a) to communicate achievement to students and parents; (b) to provide information for students for the purpose of completing self-evaluations; (c) to identify students for specific educational programs; (d) to provide incentives for students to learn; (e) to evaluate the level of effectiveness of instructional programs; and (f) if necessary, to provide evidence in
support of students’ lack of effort, lack of responsibility, or decline in behavior expectations.

McMillan (2008), O’Connor (2007), Winger (2005) and Wormeli (2006) agree with the first five categories from Guskey’s research. However, the last statement concerning non-academic factors is not part of what McMillan, O’Connor, Winger and Wormeli would describe as a purpose of grading since they believe that student behavior, effort or responsibility should not be reflected in student grades.

On another note, in support of Guskey (2004) and McMillan (2008), Reeves (2008) and Brookhart (2009) also believe that the purpose of grading is to provide accurate, specific and timely feedback to increase student performance, previously described as formative assessment. When students are provided meaningful feedback in terms of comments and not solely scores, Reeves believes that students use the comments and suggestions to improve their work, thereby making a connection to their performance and assigned grade. This allows them opportunities to improve their work and reach success, as evidenced by their learning.

O’Connor (2007) believes that a shared vision of the purpose of grades among teachers within schools and within a school district is essential. However, Kain (1996) reports that teachers fail to communicate effectively within the same school concerning grades, leaving teachers to determine their own grading practices, which usually mirror those of the school’s norms. This is supported by Guskey’s (2004) belief that teachers often reflect on their experience as a student to determine grading practices as a classroom teacher.
Researchers agree that one of the most common purposes of grading is to communicate student achievement with students and their parents (Brookhart, 2009; Guskey, 2004; McMillan, 2008; O’Connor, 2007; Wormeli, 2006). However, since it has been established that some grades reflect factors other than achievement or performance grades, Cross and Frary (1999) and Winger (2005) recommend a form of supplementary communication to report performance and progress in relationship to non-academic factors. A separate communiqué would describe work habits, responsibility and attitudes. Winger believes that a separate report could be included with a standard report card, both communicating clear messages about the student’s ability to demonstrate mastery of content and the student’s ability to follow expected work habits and responsibility – neither distorting the value of either report.

O’Connor (2007) defines achievement as performance measured against accepted standards and learning outcomes. Wormeli (2006) supports the idea that grades are intended to provide an accurate indicator of achievement factors, such as a student’s mastery of learning standards. In support of O’Connor and Wormeli, Dyrness and Dyrness (2008) agree that academic factors should be the sole component to account for an accurate measure of student achievement. Placing emphasis on student’s self-esteem, most especially in the middle school years, Dyrness and Dyrness identify the challenge of separating academic merit from behavior incentives. They believe that some teachers mistakenly use grades as a motivator with middle school students and should find other ways to motivate students to learn. Suggestions include using course materials that relate to students’ lives; allowing students to have some degree of control over their learning;
assign challenging, but achievable tasks; arouse students’ curiosity about the topics, and design projects that allow students to share new knowledge with others.

Brookhart (2009) believes that since grades serve a variety of conflicting purposes (see Table 1 for a comparison of grading purposes), the issue of conveying grades accurately and objectively remains controversial, albeit crucial components of the educational processes. Marzano (2000) was surprised with how educators defined academic achievement when they categorized achievement into three categories of information and skill, which included subject-specific content, thinking and reasoning skills, and general communication skills. Grades must communicate meaningful and accurate information about student performance. However, as there are differences in defining the purposes of grades, differences in grading practices also exist.

**Grading Practices**

One of the primary responsibilities of classroom teachers is to report grades based on student learning. In classrooms across the country, students learn a variety of content, and teachers are required to assess students’ knowledge of this content and summarize the learning into a letter or numerical grade (Allen, 2005). Simply put, there is much variability in grading practices which offers little consistency across schools and within classrooms, even when schools and teachers have adopted the same grading guidelines (McMillan, 2008). In support of McMillan, Dockery (1995) also believes that grading practices are often arbitrary and vary from teacher to teacher.

During the course of teacher preparation classes, novice teachers have limited knowledge and training in grading methods and the effectiveness of grading practices (Guskey, 2004). In teacher preparation courses, emphasis is placed on the design and
delivery of instruction, but little emphasis is placed on developing appropriate measures of assessment and contributing factors to consider when assigning grades, which supports the findings of Brookhart (1994) and Stiggins and Bridgeford (1985). Teachers are not equipped in developing assessments based on valid measurement standards.

Since grading student assignments with fairness is essential in determining accurate grades, Ebel and Frisbie (1986) attempt to determine the source of the controversy by identifying three critical factors. The first factor is that a technical challenge exists in accurately measuring achievement – suggesting that teacher preparation courses need to emphasize measurement theory, supporting the findings of Brookhart (1994) and Stiggins and Bridgeford (1985). Secondly, variations in educational philosophies among educators exist. This supports Guskey’s (2004) belief that teachers have deep-rooted philosophies about grading which are tied to their own experiences as a student. The third factor is that teachers face personal conflict when acting as an advocate and judge of their students’ performance.

In an effort to determine why teachers have difficulty assigning grades, Tombari and Borich (1999) narrow the reasons to three. The first supposition is that practices vary from teacher to teacher when assigning grades which are tied to academic achievement and non-academic factors, such as responsibility and attitude. Second, teachers have difficulty reporting one single grade for students when the grade represents multiple components, which may be an inaccurate or invalid measure of the student’s performance. Finally, teachers are inadequately trained in the areas of assessment and grading, supporting the findings of Guskey (2004), Brookhart (1994) and Stiggins and Bridgeford (1985).
Table 1

*A Comparison of the Purposes of Grading from the Literature to Date*

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Ability matches performance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Accurate measure of student</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Mastery</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Communicate achievement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Students complete self evaluations</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Provide meaningful feedback</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Identify students for educational programs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Incentives for student learning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Evaluate instructional programs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Provide evidence of non-academic factors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
</tbody>
</table>
Brookhart (1993) examined the current grading practices among 84 classroom teachers to investigate teachers’ interpretation of grades and the value judgments they made when assigning grades. The findings of this study were consistent with her 1991 finding regarding the *hodgepodge* of grading – meaning that grades often consist of an assortment of variables, which include achievement as well as effort and attitude. The phrase *hodgepodge grading* has been widely used in reference to studies involving grading practices. Contrary to the study of Stiggins, Frisbie and Griswold (1989), Brookhart believes that the reasons teachers are not in alignment with measurement experts goes beyond the fact that teachers are not aware of the recommendations, but because teachers strongly believe that effort is a relevant factor in grading, as well as achievement. Brookhart concluded that teachers considered students’ level of ability and effort when assigning final grades. McMillan (2001) and Duncan and Noonan (2007) use the term enabling behaviors when describing teachers’ use of non-academic factors such as effort, ability, attention, and work habits.

McMillan and Nash (2000) found that among high school teachers, effort and participation, along with achievement, were factors in determining grades. Furthermore, Bursuck, Polloway, Plante, Epstein, Jayanthi, and McConeghy (1996) found that teachers used a variety of non-academic factors in determining grades in special education classes, which demonstrates the pervasiveness of grading practices in all classrooms – special education and regular education. These factors included effort, notebook completion, attendance, class participation, and student organization. In Zoeckler’s (2007) study of high school English teachers, he found that teachers were influenced by perceptions of student effort and concern for the moral development of the student, meaning that the
conclusions of the study determined that equity, accuracy and honesty were questioned in
terms of teachers’ grading practices.

In 2004, a change in grading policy made headlines when a school proclaimed
that student effort would become a required substantial component of all students’
grades. Including this component became controversial, since effort would increase the
overall grades of hardworking students who achieve poorly, but students who achieve
extremely well requiring little effort would suffer a grade loss making their grades an
inaccurate measure of their level of academic achievement (Sadler, 2010).

Randall and Engelhard (2009) examined grading practices among 516 public
school teachers. Of notable interest in the study was the interpretation of how teachers
graded student work. The results of the study indicated that teachers generally assign
grades based on performance on day-to-day assessments, such as homework, quizzes, and
tests, yet the students’ final grades for reporting purposes factored in other non-academic
criteria such as effort, ability, and behavior.

According to McMillan (2008), many components are used in grading: teacher’s
intuition, subjective evaluations, and teacher’s values, beliefs, and philosophies.
Although Wormeli (2006) agrees that the use of non-academic factors is a common
practice, he cautions teachers in using subjective components. Wormeli believes that
when teachers make arbitrary decisions about students’ grades, the factors of effort and
responsibility usually do more harm than good, and it clearly sends the message to the
students that failure in responsibility affects student achievement. Wormeli proposes that
grading non-academic factors is a misuse in reporting student achievement, which can
lead to students’ perception of failure.
Contrary to Wormeli’s (2006) beliefs, the findings in Brookhart’s (1994) and Cross and Frary’s (1999) studies concluded that teachers and students found the practice of including achievement and non-related achievement factors were reasonable when determining grades, described as *hodgepodge grading* (Brookhart, 1993).

According to McMillan (2001), the findings of his study replicate findings of previous studies that demonstrate that secondary teachers use a variety of practices in determining student grades. McMillan found that the four primary categories used by teachers to determine grades are: (a) academic achievement; (b) academic enablers, such as effort, ability, improvement and participation; (c) use of external benchmarks as a standard that is defined by external accountability tests; and (d) use of extra credit and borderline cases (cases in which a student’s composite achievement average is relatively close to a grade cutoff point).

Using a different perspective in terms of grading practices, Holmes and Smith (2003), conducted a survey investigating students’ opinions about how teachers grade their assignments. The results of the investigation led to two major conclusions. The first is that students reported that the issue of grade fairness was a concern because the grades appeared inconsistent with other students’ similar work that earned different grades. Second, students reported that teachers do not provide sufficient feedback about an explanation of the grade and/or how to improve the grade. Holmes and Smith suggest that teachers set clear assignment objectives and use rubrics when grading to reduce negative student comments and opinions.

In the current era of accountability and high-stakes testing, which has promoted standardized curriculums and standardized assessments, the debate over standardized
grading practices still remains an issue. O’Connor (2007) asserts that traditional grading practices continue even with the emphasis placed on standards-based teaching, learning styles, and parent communication. He believes that traditional grading practices result in ineffective communication concerning students’ level of mastery and misrepresent students’ level of learning and achievement. Measurement specialists agree that academic factors such as student achievement should be the only basis for grades, and those non-academic factors, such as ability, effort, and behavior should not be considered (Cross & Frary, 1999; Stiggins, Frisbie and Griswold, 1989).

**Influence of Electronic Grading Systems**

In the last two decades, many American schools have immersed themselves in technology. Mostly, schools have equipped the classrooms with computers for student accessibility, and some school divisions have adopted technology guidelines for individual student laptop use. The technology shift in education has now turned its attention to maximizing technology for administrative and classroom management potential (Tetreault, 2005). On the forefront of this shift is the use of electronic grading systems. Some school divisions are moving to paperless reporting methods. In other words, electronic grading systems are on the rise because these programs are designed to build databases of student performance by collecting and recording information about student performance and progress. According to Tetreault, the electronic grading systems have the potential to serve as an essential building block of restructuring school management and administration with technology.

There are many advantages of electronic grading systems. They are created to compile grades accurately, compute averages, weight scores according to the teacher’s
instructions, flag students with designated characteristics, and print reports for individuals or groups. In addition, students and parents have available access to the electronic grading systems at all times. This feature supports communication between the school and home, and students are taking responsibility of their own learning by having daily access to their grades (Vockell and Fiore, 1993).

However, if teachers within the same school have different grading practices, using electronic grading systems may prove to be problematic. For the electronic grading system to compute grades, teachers must first input data such as how to record missing work, how zeroes will affect the average, and how different assignments are weighted relative to the overall average. This may become challenging for schools and teachers who have not developed common grading practices because it will change the composition of what grades represent. In addition, the electronic grading system lessens teachers’ objectivity and is often inflexible. Mathematical calculations are precise and correct, but numerical precision is not the same as fairness, honesty, or truth. Since the grading system simply averages numbers, teachers who use non-academic factors while computing an overall grade average, such as effort, organization, and improvement, may find grading systems to be too rigid (Guskey, 2002). Moreover, some electronic grading systems may not be designed to include categories for non-academic factors.

Vockell and Fiore (1993), identify this as a limitation in electronic grading systems – that the grading scales can not reflect other conditions that a teacher may take into consideration that could affect a grade. Although one of the advantages of using an electronic grading system is the instant communication with students and parents, Vockell and Fiore caution the use of the electronic grading system to do all of the
communicating, thereby, removing the personal relationship that teachers establish with the families.

An electronic grading system known, as TeacherEase, piloted a study in 2003 to determine if student performance would improve by using the system. Results of the study concluded that students demonstrated a significant advancement in both grades and homework completion because students were motivated to improve their performance when they accessed their grades on a regular basis. TeacherEase also promotes its software as a grading system that supports standards-based grades. The study did not include a component for teacher feedback concerning teacher efficiency and productivity and how an electronic grading system affects teachers’ grading practices (Common Goal Systems, Inc., 2010).

A small-scale pilot study of the electronic grading system called Excelsior Software was conducted in 2002 within one school division. The purpose of this study was to determine if teachers and administrators found that implementing new technology to assist with administrative and classroom tasks impacted their routines. This study did not investigate teachers’ grading practices or how an electronic grading system may change teachers’ methods in calculating grades (Tetreault, 2005). With the rise of electronic grading systems, future studies are necessary to answer critical questions about teachers’ productivity, efficiency, and implications for grading practices.

Although Guskey (2002) strongly believes that electronic grading systems are useful tools, he cautions teachers with relying solely on the system of reporting grades because teachers must report grades that reflect the most accurate and fairest description of student performance and achievement. Grading requires careful planning, thoughtful
judgment, clear purpose, and communication skills – qualities that electronic grading systems do not include.

**Standards-Based Grading**

In an era of student accountability and high-stakes testing, standards-based grading is becoming an important component in educational research. In recent research, it is noted that standards-based grading, which involves measuring students’ proficiency on well-defined course objectives, is gaining popularity (Tomlinson and McTighe, 2006). Standards-based grading provides more reliable information that measures all students fairly on comparable scales, as opposed to traditional grading practice that provides a single letter grade that reflects student achievement on combined standards. Using a standards-based reporting system is a more accurate way to inform students and parents about specific areas of proficiency as well as areas of challenge.

Traditional grading systems are often subjective. The final grade at the end of a course does not give students or parents an adequate picture of progress on all standards. On the other hand, standards-based grading systems provide formative-type data and feedback to students about a set number of related standards. As a proponent of standards-based grading, Scriffiny (2008) believes that using performance assessment increases quality work from the students and helps teachers adjust their instruction. Using standards-based rubrics provides an objective means of assessing student performance. The rubrics define specific criteria in which the students are assessed, and meaningful feedback provides the areas of strengths and weaknesses for each content standard in which the students excel or need additional supports. To employ standards-based grading requires time to develop and a change in traditional practices. Scriffiny
explains that students take greater ownership of their rights and responsibilities as a student, and greater dialogue takes place in the form of discussion and reflection.

**Grading Policies**

Grading policies as well as grading practices differ. However, it is the umbrella of grading policies in which grading practices are created. Carifo and Carey (2009) recognize that grading policies are often controversial, but believe that a critical step in employing equitable grading practices is for each school to first examine their grading policies because many schools lack a uniform grading policy, resulting in variations in student assessment from teacher to teacher. Sometimes the grading policies and their inherent purposes are often at odds with another. Realizing that schools have an educational responsibility to communicate progress to students, an initial examination of policy should reflect the school’s purpose of grades. Research supports that grades should provide and reflect fair and accurate student accomplishments and achievements based on standards and learning goals. Once schools have identified the school’s purpose of grades, Carifo and Carey make two suggestions for policy review and change. The first is that changes in grading policy should be made in the framework of the school’s philosophy. Second, grading policies benefit from certain shared traits, and the practices should be simple, straightforward, and easily understood by all stakeholders, most importantly, students and parents.

Guskey (2006) suggests that teachers should base their grading practices and policies on criteria that are clear, equitable, and unprejudiced. Often within the same school system, grading policies do not mirror one another. Teachers and administrators
are left to design grading practices based on the school board policies, which are often vague (Austin and McCann, 1992).

Austin and McCann (1992) found that school board policies may or may not include information about: (a) grading policies in reference to the purpose of grades; (b) the criteria for grading; (c) the number of marking periods within a school year; (d) the grading system used by the school; (e) the weight of final exams; (f) the definition of passing; (g) the calculation of grade point averages and class rank; and (h) the guidance for teachers to calculate grades.

According to Wiggins (1988), teachers may believe that students’ grades are based on their work and achievement, but criteria used by teachers are usually misunderstood by the students, fully supporting the idea that the purpose of an assignment and how it will be graded is important to communicate to students (Brookhart, 2009; Guskey, 2004; McMillan, 2008). Sometimes thought as mysterious, the grading practices teachers employ lack common criteria or standardization due to lack of consistency within school policy (Austin and McCann, 1992).

One noteworthy grading policy that is being debated within schools and school divisions is assigning minimum grades to address failure rates. Carifo and Carey (2009) report that the practice of assigning minimum grades to low-performing students serves as a motivator for student performance. Proponents of this policy claim that students become hopeful, confident, and optimistic about passing a course or grade, whereas, opponents of the policy believe that using minimum grades softens competency requirements and contributes to overall grade inflation.
Another recent grading policy change has taken affect in Fairfax County, Virginia when the school division advocated for major revisions of the grading policies to reflect a 10-point grading scale as opposed to a six-point grading scale that is widely used in public education. The proposed 10-point grading scale is used to assign letter grades based on a 10-point range of student grades. The program initiative, FAIRGRADE, was adopted in January 2008 to reflect change in how students’ grades are computed and weighted in order to be competitive with other students for college admissions, merit-based scholarships, and admissions to honors programs. In the research to support FAIRGRADE, school officials researched 44 of the top 100 Gold Medal Schools from U. S. News and World Reports that reflected geographic diversity in grading scale usage. Of the 44 schools, 33 of the schools already employ a 10-point grading scale. FAIRGRADE supports policy change to reflect a 10-point grading scale with pluses and minuses, and increased weights for Honors and AP/IB classes. Fairfax students outperform students from the state of Virginia and across the nation in SAT scores, yet their grade-point-averages do not reflect similar achievement because Fairfax grading policies have been in place since 1962. This is a clear example of how a school division reviewed and revised their grading policies to reflect needed change in terms of student performance, as measured by standards from other school systems and college admission programs (Fairgrade, 2008).

Good Practice and Poor Practice in Measurement, Assessment, and Grading to Reflect Student Achievement and Accountability

Good practice can be described as using evaluation tools that are steeped in measurement standards, assess the intended purpose by employing formative assessment
strategies before summative assessment takes place, and by utilizing grading practices that consider academic factors that relate to student performance. Other factors that relate to student work habits, attitudes, and behaviors are considered non-academic factors that measure student accountability in terms of behaviors, not performance. Table 2 presents a comparison of what experts say are good practices compared to what research shows actually takes place.
<table>
<thead>
<tr>
<th>Experts</th>
<th>Reality</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Good Practice</strong></td>
<td></td>
</tr>
<tr>
<td>Use only achievement grades</td>
<td>Achievement grades and non-academic behaviors, attitudes and work habits are used</td>
</tr>
<tr>
<td>Communicate with students and parents</td>
<td>Communication takes place with students and parents</td>
</tr>
<tr>
<td>Provide meaningful feedback using comments and suggestions</td>
<td>Feedback is provided, but sometimes lacks meaningful comments</td>
</tr>
<tr>
<td>Use formative assessments as much as possible while students are learning new material</td>
<td>Summative assessments are widely used, mirroring those of standardized testing</td>
</tr>
<tr>
<td><strong>Poor Practice</strong></td>
<td></td>
</tr>
<tr>
<td>Use a combination of achievement grades and non-academic behaviors to report student grades</td>
<td>Use a combination of achievement grades and non-academic behaviors to report student grades</td>
</tr>
<tr>
<td>Primarily use summative assessments</td>
<td>Primarily use summative assessments</td>
</tr>
<tr>
<td>Do not communicate with students and parents about student progress and performance</td>
<td>Do not communicate with students and parents about student progress and performance</td>
</tr>
<tr>
<td>Grades are not based on the expected performance of students</td>
<td>Grades are not based on the expected performance of students</td>
</tr>
</tbody>
</table>

Summary

The practice of assigning grades is complex. Typically, schools within the same school district do not employ the same grading practices; neither do teachers within the same school. Teachers tend to develop their own grading practices based on professional norms and previous experiences. Allen (2005) believes that grading practices need to improve to be more accurate and educationally meaningful. He proposes two major changes. The first is to determine student grades based solely on achievement, without factoring in non-academic behaviors. If the non-academic behaviors are important to communicate, then the behaviors need to be reported separately. This method of communication is supported by Cross and Frary (1999) as well.

Secondly, teachers need professional development in terms of creating effective measures of assessment and in providing meaningful grading practices that measure student achievement. Allen (2005) reports that one of the main reasons teachers have difficulty determining appropriate grading practices used to measure student performance and achievement is due to the lack of professional development in the areas of measurement — how to create tests that measure the intended purpose, and how to subsequently grade the measurement. Throughout the research, in terms of aligning measurement, assessment, and grading practices, one of the common themes is the lack of professional development for teachers (Brookhart, 1994; Ebel and Frisbie, 1986; Stiggins and Bridgeford, 1985; Tombari and Borich, 1999; Guskey 2004).

When teachers are not aligned in terms of developing assessments that meet measurement standards and assigning grades based on good practice and policy, the issues of reliability and validity of student grades are questioned (Stiggins, 1999; Guskey,
2006; Gallagher, 1998). Allen (2005) and Stiggins and Bridgeford (1985) recommend that coursework for pre-service teachers and professional development for current teachers take place to facilitate alignment. This will eliminate the disconnect among the three factors of evaluation (measurement, assessment, and grading) and truly measure student achievement in accurate terms.
Chapter III

METHODOLOGY

This chapter presents the methodology of the study. The design of the study will be explained, and the setting and participant selection will be described. Data collection and data analysis components will be detailed. The chapter concludes with the Institutional Review Board process.

A qualitative study of ten middle school language arts teachers was conducted to determine how and why they use academic and non-academic factors in grading. Language arts teachers were selected to complete the qualitative interviews to keep the comparison of teachers’ grading practices within one core subject. The core subject of language arts was chosen because there is much teacher subjectivity and opinions associated with grading of language arts assignments, such as writing pieces, journals, research papers, and essays. Language arts teachers focus on different criteria for different assignments. The choice to use only one department was made in order to gather data used for comparing and contrasting within the same content. In addition, studying one area in depth will provide different and more meaningful perspectives from one area of study that is rich in subjectivity in terms of grading written and oral responses in the forms of presentations, essays and term papers, contrary to concrete and precise answers found in most mathematics and science works.
Prior approval from the school district was sought and granted. Participation was strictly voluntary. During the qualitative interviews, selecting the site was important (McMillan and Schumacher, 2006). Selecting the right site that is accessible, suitable, and available for the participants is a critical component of qualitative research. The qualitative interviews took place in a quiet setting that limited interruptions and noise. Research was conducted in a natural setting so the researcher could discover the reasons behind teachers’ grading practices in real-time (Guba and Lincoln, 1988).

**Research Questions**

Four research questions emerged that focused on the grading practices teachers employ and why these practices were used. These questions served as a guide for the researcher and were adjusted during the study to reflect the anticipated research that would be found during data collection. The four questions were:

1. What influences teachers’ grading practices?
2. What academic factors were considered, and what beliefs and/or values were used to determine those factors?
3. What non-academic factors were considered, and what beliefs and/or values were used to determine those factors?
4. What gaps exist between teachers’ grading practices and recommendations made by measurement experts?

**Design**

The primary data collection involved interviews with ten teachers to determine their reasons for using both academic and non-academic identified factors when assigning student grades. In recent research studies, many researchers have used
quantitative data to focus on the types of grading practices teachers employ. In this study, a phenomenological design was used to describe and analyze the reasons behind the teachers’ grading decisions to make sense of their individual and collective choices involving grading practices. Since all participants share similar experiences in assigning student grades, a phenomenological study was appropriate to focus on one shared phenomenon to gain an in-depth understanding (McMillan and Schumacher, 2006).

According to Creswell (2007), a phenomenological design is used to develop a composite description of what the participants have experienced, and how they have experienced it.

Appropriate for a phenomenological study are in-depth interviews that produce informative data in terms of the research questions. Generally, the broad questions focus on what the participants have experienced and what factors of influence affected the experience (Creswell, 2007). An interview guide was used (see Appendix A) during the audio-taped interview, and participants were asked questions which focused on what considerations were used in grading, and how the factors relate to measurement theory. Questions were open-ended, which allowed for more in-depth responses, and at times, the participants’ answers elicited more probing questions. The average duration of the interviews was 45 minutes, and only one interview per person was necessary.

**Rationale for Design**

In interview research, a qualitative design is appropriate to understand the situation from the perspectives of the participants (McMillan and Schumacher, 2006). A qualitative approach was appropriate for this study because the qualitative data collection with a small sample of teachers focused on an inquiry of understanding a social or human problem which involved detailed views from the participants (Sogunro, 2001; Creswell
This orientation is reflected in this research in studying teachers’ choice in different grading practices. Frankel and Devers (2000) describe qualitative research as a process that is non-linear and non-sequential, where data collection and analysis are often simultaneous. The goal of qualitative research is to better understand cultural phenomena, and consists of text rather than numbers as data. Qualitative analyses tend to be ongoing and iterative, meaning that data are analyzed as collected, which may affect further data collection efforts. Therefore, at the end of each interview, the researcher will consider the responses and modify the next set of questions, collecting and analyzing data during each successive interview.

**Participant Selection and Setting**

A purposeful sampling method was used in this study (McMillan and Schumacher, 2006). This approach was not a random sampling, but a sampling of participants who are knowledgeable and experienced with grading practices. Participants from a purposeful sampling method tend to be volunteers. Bias of this type of sampling includes a concern that volunteers tend to be better educated, more intelligent, more extroverted, and more sociable, which restricts the generalizability of the results (McMillan and Schumacher).

Appropriate for this study, maximum variation sampling was used to select subjects on the basis of characteristics of the population. According to McMillan and Schumacher (2006), certain quotas are met when the sample represents the intended population, based on the characteristics. The characteristics of the sample for this study included language arts teachers at the middle school level within one mid-size central Virginia suburban school division, representing four middle schools. The school division
serves approximately 18,000 students, and the middle school population is roughly 4,500 students and employs approximately 400 teachers. Only language arts teachers were asked to participate to keep the comparisons of teachers’ grading practices within one core subject.

According to McMillan and Schumacher (2006), using a maximum variation sampling will assist with obtaining maximum differences of perceptions about the topic from well-informed participants. The process of sampling began with an email to all language arts teachers at all four middle schools (approximately 50 teachers), explaining the study and then asking for volunteers to participate in qualitative interviews concerning the grading practices they use. For the volunteer list, two questions were asked. The first question was a general one about the teacher’s current practice of calculating the percent of a student’s grade determined by academic performances as contrasted by the percent determined by non-academic factors. To assist with which factors are considered academic and non-academic, definitions of terms were included in the email. The second question asked volunteers about their availability for an interview. From the answers to the first question, the scores in the highest and lowest tenth of volunteering teachers concerning their inclusion of non-academic factors in grading produced the list of teachers to be interviewed. This information determined the maximum differences of teachers’ grading practices based on their percentages of those teachers who grade primarily on academic factors only and those teachers who use a mixture of academic and non-academic factors to assign a grade. Qualitative interviews were conducted with the available sample until a saturation of data existed.
Pilot Test

A pilot test was conducted on the qualitative interview questions during the summer months of 2010 to identify and measure the interview content based on the specifications of the study. Three current middle school language arts teachers were asked to review and make anecdotal comments about the (a) clarity of the interview questions; (b) bias of the interview questions; and (c) overall flow of the interview questions based on the topic of teacher grading practices. After explaining the purpose of the pilot test, along with the purpose of the study, each interview took about 45 minutes to complete. All three teachers were comfortable with answering questions based on the grading practices they employ, and were interested in learning more about the recommendations made by measurement experts in terms of grading. At no time were they uncomfortable in explaining their decisions and they were open-minded in sharing their thoughts on how to improve the questions. Their responses were used to revise the interview protocol to ensure that the understanding of the questions matched the stated purpose of the instrument.

Instrument Revision

The pilot test participants provided meaningful feedback in terms of the clarity of the questions in relationship to the stated purpose of the study. Suggestions were made to revise the interview protocol to: (a) include more specifics about the definition of terms in contextual reference; (b) change the order of the interview questions in terms of what academic and non-academic factors teachers use; (c) include questions about the criteria teachers use in creating assessments; and (d) include additional questions about teacher training in the area of grading practices.
Data Collection

Once the University’s IRB approval was obtained, the school division approval for teacher participation was also sought to gain approval to conduct qualitative interviews with the division’s school teachers. The sample population was then established, and the qualitative in-depth interviews took place. In qualitative interviews, an appropriate data collection strategy is humanistic, as in personal interviews (Creswell, 2007). During audio-taped interviews, participants answered questions from an interview guide (see Appendix A). The questions were open-ended, which allowed for more in-depth responses, and led to more probing questions.

In qualitative research, it is important to map the field (McMillan and Schumacher, 2006). Mapping the field is also known as gaining entry into the site or establishing rapport with the participants. Establishing rapport with the teachers who participated in the qualitative interviews assisted in their comfort level throughout the interview (Creswell).

The participants were not rushed to answer due to the reflective nature of the questions. All teacher participants gave consent to participate in the study (see Appendix C), and the participants were assured of their confidential responses. In the interview guide some questions were related to demographic information, yet the majority of the questions were open-ended, relating to the topic of grading practices. At times, the order of the questions varied, depending on the flow of the conversation with each participant (Creswell, 2007).

The interviews were conducted at a location convenient for all parties, with emphasis placed on a quiet setting with few or no interruptions. A transcriptionist
transcribed each interview within two days of each interview. Each participant received a copy of the interview to review and request changes, ensuring the validity of their responses. This process is known as member checking (McMillan and Schumacher, 2006).

All interviews were kept as separate word document files in Microsoft Word, and the files have been saved in four different places – two removable disks, one laptop computer and one hard drive on a personal computer.

Data Analysis

Once the interviews were transcribed, a general analysis of each interview took place, looking for emerging themes during the discovery phase. The analyses of the transcribed interviews, from coding to comparing and contrasting, were completed with the assistance of a qualitative research software program, NVivo.

Deconstructing the interviews was the next step. During this process, codes were developed to identify recurring ideas and major themes. The literature supported that codes would emerge based on the reasons teachers assign grades using academic and non-academic factors, as well as the limitations teachers experience in terms of staff development in the areas of developing assessments and implementing fair grading practices (Allen, 2005; Ebel and Frisbie, 1986; Guskey 2004). Results from the qualitative data analyses were shared in narrative form, with emphasis placed on the comparison and contrast of the participants’ responses (McMillan and Schumacher, 2006), which produced substantive and conclusive themes and subthemes.

The judgment that qualitative analysis is trustworthy is based on the rigor of the process used for collecting and analyzing data. Three techniques were used to confirm
the interpretations of the data to establish internal validity, also known as credibility. Participant language was employed by using direct quotes within the report to illustrate the participants’ meaning and to ensure validity. Secondly, participants reviewed the findings of the researcher’s syntheses to ensure accuracy of the representation. Finally, negative cases or discrepant data were reported in findings that contradicted the emerging patterns. These are known as outliers of the findings. To determine dependability, also known as reliability, it was determined that the data measured the original intent of the research (McMillan and Schumacher, 2006: Creswell, 2007). Reliable data are consistent data that emerged from the study’s participants.

It was expected that teachers will not give less than honest answers, even though the researcher is a school administrator within the same school division, which was affirmed during the pilot test. Two strategies used in qualitative research to eliminate researcher bias were used. First, feedback from key informants was sought. Known as member checking, this strategy strengthens the credibility of qualitative analysis. Employing member checking involved asking key informants to read the researcher’s report to verify that the information was accurate. Secondly, soliciting feedback from external expert reviewers, called peer review, also strengthens credibility. Key informants and external qualitative methodology experts question the analyses; push the researcher to better clarify their thinking, and verify the meanings and interpretations of the results (Strauss and Corbin, 1990).

**Delimitations**

In research studies, it is the expectation that the findings of research are generalized and have implications for further research. Typically, quantitative
researchers use large, random samples to enhance the generalizability of statistical findings. However, in qualitative studies, generalizability is restricted due to low samples, making it not possible to satisfy statistical generalizing. Appropriate for qualitative studies, transferability will be used to judge the extent to which the findings can be applied to other contexts. The results of this study may be transferable if the range of contexts and conditions of this study is realistically replicated in a similar inquiry. Strategies used to attain transferability include thick descriptions and purposeful sampling. When thick descriptions are used to describe the contexts, conditions, samples, and categories of the study, future researchers may judge the appropriateness of applying the findings to other settings.

Guba and Lincoln (1988) believe that transferability is a process also performed by researchers, by which they are able to infer that the results of the research would be similar in their own situation. The degree of transferability is a direct function of the similarity between the two contexts of future studies, since qualitative research is dependent of a particular time, place, and population. Therefore, in this study, delimitations exist due to the sample. The study is delimited to language arts middle school teachers, grades six through eight, and teachers in central Virginia.

**Institutional Review Board**

The purposes of the Institutional Review Board (IRB) are threefold: to review research activities involving human subjects, to ensure that ethical standards in working with human subjects have been established, and to ensure that research activities are in compliance with federal, state, and local policy – all to ensure the protection of human subjects in research. All proposed research studies must be submitted and reviewed for
approval before the study begins. For this study, the IRB proposal included sections
detailing the significance and purpose of the study, literature background on the subject
of grading practices, and a detailed account of the proposed methodology to conduct the
study.

Once the IRB proposal was approved, obtaining informed consent from
participants was needed. Elements considered when obtaining consent included (a)
informing the participant who would conduct the study; (b) giving reasons why the
participant was asked to participate; (c) explaining the time commitment involved; (d)
explaining benefits to participating in the study; (e) explaining potential risks and how
they would be managed; (f) explaining the purpose of the study; (g) explaining that
participation was voluntary; (h) providing a written copy of the informed consent form;
and (i) explaining that maintaining the participants’ confidentiality would be critical
throughout the study.
Chapter IV

FINDINGS

In this chapter, the findings or results of the ten qualitative interviews are explained in narrative form. The findings were categorized into four major themes and subthemes which emerged within each major theme. Throughout the chapter, comparisons and contrasts naturally materialize among the participants.

Participants

The following data describes the beliefs and grading practices of a sample of ten middle school language arts teachers. All ten middle school language arts teachers are from the same mid-size suburban Virginia school division. The school division is one that has grown in reputation and in enrollment over the past twenty years largely due to an increase in student achievement in standardized test scores and overall rigorous curriculum. The school division serves approximately 19,000 students in grades Kindergarten through twelfth. The ethnicity of the student population within the school division is predominantly white, making up 84% of the population; 9.5% of the students are African-American, and the remaining 6.5% of the students are Hispanic, Asian, or American Indian. Approximately 14% of the students receive special education services, and 12% of the students receive gifted and talented services. The middle school population is approximately 5,900 students.
After receiving approval from the school division and each middle school principal, the language art teachers were asked to participate in the study through an electronic letter, outlining the basic concepts of the study, and what would be involved if they chose to participate. All participation was voluntary, and to ensure confidentiality, the participants are identified by pseudonyms, not their proper names. The participants in this study are all female, with an average age of 41, and have an average of 15 years teaching experience.

Table 3

<table>
<thead>
<tr>
<th>Gender</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age</th>
<th>21-30</th>
<th>31-40</th>
<th>41-50</th>
<th>51-60</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Years of Teaching Experience</th>
<th>1-10</th>
<th>11-20</th>
<th>21-30</th>
<th>31-40</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

All participants currently taught at the middle school level. The first participant, Lee, has been teaching for nine years, seven of which were at the elementary grade levels of fourth and fifth grade, while the last two have been at the middle school level teaching seventh grade. Lee earned her Bachelor’s Degree in Liberal Studies and recently earned a Master’s Degree in Administration and Supervision. In the future, Lee plans to pursue a career change in school administration.

Edna, who serves as the second participant, has the most teaching experience from the sample. All 36 years of teaching experience have been in the area of language
Edna has taught arts in middle school or English in high school. She has taught at her current school for ten years, and she has been nominated and received many awards for her effective instruction. Most notable is her R.E.B. Award for Teaching Excellence from the Community Foundation. Edna earned her master’s degree in liberal arts, and her undergraduate degrees are in English and fine arts.

Participant number three is Tyler, who has been teaching for six years. She is the second youngest in the sample, and all six years of teaching experience have been with middle school students, teaching language arts. Tyler earned a Bachelor of Arts degree in English from James Madison University. She continued her education by earning a Master’s of Church Ministries in Christian Education from Duke University.

Cooper, serving as participant number four, is the youngest teacher of the sample. She has been teaching for four years, and all of her experience is at her current school, teaching language arts. Cooper’s Bachelor’s Degree is in secondary education, integrated language arts, and she is currently working on a Master’s Degree in Counselor Education through Virginia Commonwealth University.

Mary, participant number five, has nine years of teaching experience, but they are spread over time, having taken a long break from education while staying at home with her children. Her first years of teaching were in elementary school, teaching Kindergarten and first grades. When Mary returned to teaching, she began her current assignment of teaching middle school language arts.

Serving as participant number six, Smith’s educational experience ranges from teaching elementary school to serving as a reading coach. Her current position as the reading coach in the middle school gives her an opportunity to continue with teaching
responsibilities as well as mentoring and coaching new and veteran teachers in the areas of language arts. Smith also serves as the facilitator for her school for the content area Literacy Grant Program which has a direct impact on other content classes as well. She earned her master’s degree as a reading specialist and has impacted the reading scores at her current school through collaboration in instruction and assessment.

Michelle graduated from James Madison University with a Bachelor’s Degree in English. Her first career was in journalism with a small newspaper, and after seven years decided to switch careers to education. She began teaching five years ago at her current school in the area of language arts. Michelle finds great strength in collaborating with other language arts teachers and the school’s reading coach to enhance her instruction. She is skilled in analyzing benchmark data to help her make instructional decisions.

The eighth participant is Ann. Ann began teaching right out of college sixteen years ago. She earned her bachelor’s degree in English, with a minor in education from Randolph-Macon College. All of her experience has been with middle school students, in the area of language arts. Ann serves as a teacher-leader in her current school as the department chairperson. During their monthly department meetings, Ann assists the department by analyzing student data and making informed decisions about instructional planning. She is an advocate for collaboration, and she has earned the respect of her colleagues. Ann returned to graduate school within the last three years to earn her Master’s Degree in School Administration. Similar to Lee, Ann is actively pursuing a career change in school administration.

Although using the pseudonym Fred, participant number nine is indeed a female who possesses a great sense of humor. Fred has had a range of teaching experiences,
ranging from music education to the gifted and talented resource teacher for middle
school. Fred is the oldest of the teachers from the sample, and collectively has been
teaching for 27 years, most of which have been in the area of language arts. In addition,
Fred’s bachelor’s degree is in music education, and she earned a Master’s of Education
Degree in Curriculum and Instruction. Pursuing her desire for school administration,
Fred served as a grade-level administrator for four years. She then decided to return to
the classroom full time where she believes she had the made the greatest impact with
instruction and student achievement. Since Fred enjoys new learning opportunities, she
actively engages in professional development by attending state and national conferences.

Final participant Rachel has been teaching for 20 years, all in the areas of
language arts, English, and special education. She first began teaching in high school,
and after 13 years, changed schools to teach middle school language arts. Rachel’s
undergraduate work was completed at Longwood University, and she earned a Master’s
Degree in Reading from the University of Virginia, completing the requirements for a
reading specialist certification.

Data Analysis

Ten individual audio-taped interviews took place in secure environments to
ensure privacy and confidentiality. A comfortable rapport was established between the
participant and researcher, and the participants elaborated on the open-ended interview
questions, adding honest and relevant information to their responses. During the
interviews, member checks occurred when the researcher restated or summarized the
information and then questioned the participant to determine accuracy. In addition, the
researcher wrote anecdotal notes about the overall impressions during and immediately
following each interview. Once the interview was transcribed, each participant reviewed her interview and the researcher’s interpretations to assure the accuracy of her responses. Member checks occurred at the conclusion of the study by sharing all of the findings with the participants, allowing the teachers to critically analyze the findings. The participants either confirmed or disagreed with the summaries which reflected their views, feelings and experiences.

During the qualitative interviews, questions concerning grading practices ranged from the type of factors considered when grading to conflicts they experienced once they developed their system. The first component of the analysis began with the development of the coding system constructed after careful review of the transcribed interviews, using the qualitative software program NVivo. The analysis of the interviews yielded individual text references, categorizing the information into four main themes with subthemes for each. The number of text references for each interview is listed in Table 4. The analysis of each interview continued to take place through personal reflections, constant comparisons, and the use of intercoder reliability.
Table 4  
*Code Frequency*

<table>
<thead>
<tr>
<th>Participant</th>
<th>Interview #</th>
<th>Number of Text References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lee</td>
<td>1</td>
<td>34</td>
</tr>
<tr>
<td>Edna</td>
<td>2</td>
<td>36</td>
</tr>
<tr>
<td>Tyler</td>
<td>3</td>
<td>29</td>
</tr>
<tr>
<td>Cooper</td>
<td>4</td>
<td>33</td>
</tr>
<tr>
<td>Mary</td>
<td>5</td>
<td>25</td>
</tr>
<tr>
<td>Smith</td>
<td>6</td>
<td>34</td>
</tr>
<tr>
<td>Michelle</td>
<td>7</td>
<td>35</td>
</tr>
<tr>
<td>Ann</td>
<td>8</td>
<td>35</td>
</tr>
<tr>
<td>Fred</td>
<td>9</td>
<td>36</td>
</tr>
<tr>
<td>Rachel</td>
<td>10</td>
<td>32</td>
</tr>
</tbody>
</table>

In Table 5, the four main themes and subsequent subthemes are listed. The number of text references found in the number of sources is also listed.
Table 5  
*Data Analysis Themes and Subthemes*

<table>
<thead>
<tr>
<th>Theme</th>
<th>Number of Text References</th>
<th>Number of Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purpose of grading</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Communication</td>
<td>18</td>
<td>10</td>
</tr>
<tr>
<td>Provide feedback</td>
<td>39</td>
<td>10</td>
</tr>
<tr>
<td>Student progress and mastery</td>
<td>13</td>
<td>8</td>
</tr>
<tr>
<td>Grading policies and procedures</td>
<td>15</td>
<td>10</td>
</tr>
<tr>
<td>Factors considered when grading</td>
<td>66</td>
<td>10</td>
</tr>
<tr>
<td>Influences</td>
<td>14</td>
<td>10</td>
</tr>
<tr>
<td>Academic factors</td>
<td>16</td>
<td>9</td>
</tr>
<tr>
<td>Non-Academic factors</td>
<td>17</td>
<td>10</td>
</tr>
<tr>
<td>Student behaviors</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td>Student effort</td>
<td>22</td>
<td>9</td>
</tr>
<tr>
<td>Student responsibility</td>
<td>15</td>
<td>10</td>
</tr>
<tr>
<td>Factors deemed most important</td>
<td>14</td>
<td>9</td>
</tr>
<tr>
<td>Assessment</td>
<td>49</td>
<td>10</td>
</tr>
<tr>
<td>Formative</td>
<td>39</td>
<td>10</td>
</tr>
<tr>
<td>Summative</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Professional Development</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teacher training in grading practices</td>
<td>13</td>
<td>10</td>
</tr>
<tr>
<td>Teacher conflict with own practices</td>
<td>17</td>
<td>9</td>
</tr>
<tr>
<td>Need for professional development</td>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>

**Purpose of Grading**

Grades certainly have a purpose in the American educational system. Before the participants began communicating about which factors they considered when approaching their grading systems, they first expressed their ideas about the purposes of grading and why grading is important. Although grading practices may be considered arbitrary, the purpose of grading is not a hotly-debated topic. When asked to define the
purpose of grading, all ten teacher-participants answered quickly, as if this were an area that withstood investigation. Collectively, four themes emerged when describing the purposes of grades: (a) to communicate with students and parents; (b) to provide feedback to students; (c) for students to demonstrate progress and mastery over time; and (d) to adhere to grading policies and procedures.

Communicate with Students and Parents

There were 18 separate comments that were coded purpose of grading as it relates to communication, noting only minor differences from all participants. This suggests a common and comprehensive set of beliefs that a primary purpose of grading is to communicate. All agreed that the expectations in grading should be communicated to students and parents at the beginning of the year and throughout the year as coursework and projects are assigned. The participants all found using a course syllabus at the beginning of the school year very beneficial as a form of communication with students and parents. Each teacher was expected to distribute a course syllabus, and the teachers reported that detailed information about grading expectations alleviates student and parent misunderstanding and surprises.

As Lee reported, “It’s important to be able to communicate with your students and parents, especially in middle school. They understand, so you can have a conversation with students about how you came to a grade.” Ann agreed by stating, “Communicating grades is a must, a priority because they are not my grades, they are the students’ grades. If they don’t know their scores, they don’t know how they can improve.” Fred found the question a bit humorous by stating,
Communicating grades to students and parents is very obvious to me. I would be very upset if I didn’t know what was expected of me or how it was going to be graded, because it’s a turkey shoot otherwise. You just don’t know where you should focus.

Tyler finds that communicating with students and parents is a courtesy and helps students focus on what is expected, and in her experience, she finds it helpful to give at least one month’s notice to her students about upcoming assignment and grading expectations for the current unit. Rachel agreed that, “Communicating with parents is important and emphasizes that communicating more than one time at report card or interim time is beneficial and supports the use of electronic grading systems because of the feature of instant access.”

**Provide Feedback to Students**

Also known as formative assessment, the participants believe that providing feedback to students is a core purpose of grading, which was a clear finding in the 39 comments coded in this sub-theme. The researcher noted that the participants were passionate when speaking about this topic. In fact, some of the teachers appeared surprised that anyone would suggest that providing feedback would not be a key purpose in the grading process, attesting to the fact that all participants share a common philosophy. For example, providing meaningful feedback allows Rachel to conference with students, offering comments that she finds more powerful and valuable than a number. For example, during writing exercises, Rachel makes notations in the margins about the student’s ideas and word choice. As she stated, “Students get more from that than an 85, 89, or a 73.” Similarly, Edna believes that receiving just a number or letter
grade warrants suspicion, as she recalled a time while she was a student in graduate school who asked for evidence from her professor to support the assigned grade. Edna further recalled that the grade she received from the professor appeared arbitrary because it lacked feedback, making her suspicious of the assigned grade. Therefore, Edna finds that a true purpose of grading is providing feedback to assist students understand where their strengths and challenges are, especially in writing.

In support of Edna’s beliefs, Tyler stated, “I put my most effort in our writing prompts, highlighting the students’ glows and grows – where they are doing well and what they need to work on.” Since she also believes that grades can be used as a motivator, Tyler said,

Students gauge their improvement by letters and numbers, but by becoming motivated by numbers and letters, they lose the value of improving themselves and learning. Obviously, we have to have some type of assessment to prove a child’s progress and grades; though an imperfect system, it serves that purpose.

Smith believes that when students receive meaningful feedback, it supports the reason behind the grade. Although she admits that it is very time consuming for teachers, Smith further believes that the appropriate amount of feedback is important. She has seen student work when the teacher has bled on the paper with a red pen, and she has observed the other spectrum, where little or no feedback was offered. Realizing that providing an appropriate amount of feedback may differ from one teacher to another, Smith believes that students need feedback in order to make decisions about their growth.

Michelle supports the idea of providing both a grade and feedback as a purpose of grading when she stated, “I think the students want some sort of score because they need
some sort of measurement on how they’re doing. The grade reflects whether or not they understood it.” Related to Michelle’s thoughts, Ann agrees that students find more meaning in comments coupled with a score. Ann believes that students rise to meet the challenge of improving their writing based on the comments, not necessarily just to improve the domain score.

**Students Demonstrate Progress and Mastery**

Only 13 comments were coded concerning the topic of student progress as it relates to the purpose of grading. The researcher noted that some of the participants described student progress in terms of student effort and not so much in terms of the purpose of grading. Since students are using meaningful feedback from their teachers in formative assessments, it is expected that students will demonstrate progress over time, eventually mastering the concepts. This idea of students demonstrating progress holds the students accountable for their work as well as the teachers, reported Rachel.

Concurring with Rachel’s statement, Lee reported,

The purpose of grading is to see if students have mastered the skills and concepts we have taught them over a period of time. It’s also to see the progress or growth students have made from the beginning of the year to the end.

Edna supports the same idea when she stated,

The purpose of grading is to give students, parents, and teachers a concrete indication of what has been achieved. Without guidelines and rubrics, students won’t know what is being assessed, and without grades, they will not know what they have achieved in the eyes of their instructor.

Cooper’s thoughts are in alignment; however, she also reported,
Grading has two purposes. One purpose is to assess how students are achieving in comparison to their peers. The second purpose of grading is to show student effort and attitude. Effort and attitude should factor into grades because students should be rewarded appropriately for hard work and effort, even if he or she has not actually mastered the skill yet.

Conversely, Mary does not take into account students’ level of effort or attitude when she stated, “Grades show mastery. Knowledge builds. We expect someone to achieve mastery of one concept before moving on to the next.”

Michelle looks at the end result when describing the purpose of grades in terms of mastery and student performance when she stated,

The overall purpose of grades is to give students feedback on mastery of the subject and certain skills that they’re expected to have. The student’s end grade is a reflection on the level of mastery the student attained by the end of the year.

In alignment with Michelle’s statement, Ann also believes that the purpose of grading is to “measure the students’ ability to learn the content with grades, and that grades serve as motivation for students to improve.” Additionally, Ann states that, “We, as teachers, are judged by our students’ grades. Grades are a reflection of the teachers’ hard work.”

Supporting the idea that grades are a reflection of teacher and student accountability, Fred believes that grades quantify performance by numbers, reporting data by which parents, administration, and central office can understand – simple numbers. Fred strongly believes that anecdotal notes and summary performance by student are more valuable in reporting student progress and mastery. Fred continued by
stating, “I feel like the grade should reflect how the child is performing. Learning is growth over time, and Suzie’s A may be different than Joe’s A.”

Edna defined the purpose of grading by helping students demonstrate progress over time. She explained,

Getting a bad grade is a license for them to continually do the same thing over and over again. Giving them a second chance and saying, ‘okay, you’re going to do this right until you get a better grade,’ proves the point. It’s not a one shot deal.

Cooper supports a similar idea by using writing notebooks. Using the notebooks,

I can go in and make comments and give more feedback to students prior to taking a grade. That’s the whole point in improving writing through constructive feedback. You have to show them, and then you have to let them practice on their writing. I make comments about what they focus on, giving them true feedback, which takes forever to write. The most frustrating part is when they haven’t changed a thing in their final copy.

**Grading Policies and Procedures**

The consistency of adhering to grading policies and procedures leads to a greater understanding of the purposes of grading. When asked about grading policies and procedures, all ten participants had strong beliefs about using consistent practices, yet some still found room for flexibility. For example, Rachel stated, “I don’t think grading is just black and white. I think there needs to be a gray area that allows for some flexibility.” Conversely, Lee uses consistency in her grading practices. She stated,
If you start the year out one way, you need to finish it off the same way. Factors that I include in grading need to be consistent. Such as, if I value effort as part of a grade, then I need to value effort all year long – consistency.

Edna points out another factor in terms of consistent grading practices by using an example.

When I’m grading papers, I have to only grade so many at a time because if I tried a marathon, by the time I get to the end, I’m like fine, okay, yeah. So, the consistency for me is very important. I have to be in the same mind set when I’m grading one particular class, so I’ve learned not to do all of my classes at the same time.

Tyler agreed with Edna that it’s difficult to remain consistent when grading writing samples and projects because of the subjectivity. In addition, Tyler believes, it is hard to make it consistent across the whole school because as a language arts teacher, my grading is different than a math teacher, but within a department, I think it would be good. In the end, I think we need to let the instructors be differentiated for their organizational skills and their abilities. Some teachers are point-graders and some teachers use percentage grades, so ideally it would be more consistent, but I don’t think that the goal should be for everyone to have the exact same practices.

In Cooper’s school, the language arts teachers plan together, including the development of the same rubrics, tests, and other assessments. However, even with the common planning and development of assessments, Cooper finds that grading practices vary from teacher to teacher because “teachers have different personalities, which factor into the
atmosphere of the classroom.” Realizing that her practices are different from some of her colleagues, Cooper concludes, “My focus is different, but as long as I can justify the grades, I’m okay with it.” Mary refers to the course syllabus again during this discussion and stated,

Grading policies and procedures need to be consistent – absolutely. I think it’s easier for parents when the team has a consistent approach to grading. We communicate our grading practices in the syllabus at the beginning of each year, and we refer back to it when appropriate.

Contrary to Mary’s idea, Smith believes that her grading is subjective in the area of language arts which allows for inconsistent grading, especially when she focuses on the growth and progress of the student’s work. Michelle identifies with Mary when she stated,

I try to be consistent, but then I give some students second chances because I know that they are capable of performing better, but I’m not being consistent with some of the other students because they may already have a B, even if they might be missing some of those same assignments. I’m not going out of my way to give it back to them. As a school, I think it’s hard to be consistent through the different subject areas. I recognize that it’s good to have a blanket policy that you have these certain grades for a school because it’s easier for the students and parents to keep track of that sort of thing, but I think it is hard because different classes do different things. However, in our department in our grade level, there are only three of us, and we tend to use consistency, especially in writing. We
look at student effort, and students will not be penalized because they may not be as strong a writer as others.

Fred concurs with Michelle about individualistic effort, when she reported,

The students need a very well-defined and consistent framework in which to operate. However, not everybody is on the same abstract thinking level, so I expect the best that they can do, and I try to push them beyond where they are. I leave room for subjectivity and flexibility because I’m grading them on their thinking process and not just their answers.

Ann strongly believes that the students will not be able to meet the teacher’s expectations if they don’t know what they are. “If you’re not consistent, you haven’t set the expectations for the children to succeed.”

Summary of the Purpose of Grading

Throughout the interviews with all participants, the first important finding involved the purpose of grading. During this portion of the interview, the researcher’s reflections included moments of discovery realizing that a common thread was emerging from one interview to the next. In addition, the researcher noted that the topic of why teachers assign grades was met with thorough reflection due to their comfort level with the topic in practical experience and knowledge.

With ease, the participants explained that they have always understood that grading essentially means communicating – all sharing a similar philosophy. The different modes of communicating range from providing feedback to students and parents to using systematic approaches in sharing the results of assessments. Additional beliefs
surfaced that defining the purpose of grading also includes student progress and mastery of material and adhering to grading policies and procedures.

When the researcher considered the discussions from the participants that resulted in identifying four purposes of grading, it was evident that the participants fully understood the importance of communicating results, providing feedback, and adhering to school expectations. This suggests that teachers use practical experiences to confirm their beliefs. Once the teacher-participants defined the purpose of grading, the next topic of grading factors flowed naturally in the discussion.

Factors Considered When Grading

At the beginning of each school year, teachers spend much time in reflection, preparation, and planning for the new school year. Each year is a fresh start with a new group of students, and it is imperative for teachers to capitalize on the gains they made the previous year in terms of effective planning, instruction, and assessment. One of the topics in which most teachers reflect is on grading practices. They question themselves about what they learned from the previous year to implement for the upcoming year.

Throughout the interviews, the answers varied greatly in responses to the factors teachers considered when grading. There were 66 separate comments that were coded within the interviews, making grading factors the most referenced topic during all of the interviews. Although a hot-topic with the participants, they had different philosophies about which factors were most important, and which factors should be graded versus non-graded. Much of the discussion with each participant focused on several themes: (a) influences when determining grading practices; (b) academic factors; (c) non-academic factors such as student behaviors, student effort, and student responsibility; and (d)
factors deemed most important by the teachers. Of notable interest, teachers with little experience appeared to be more stringent with their grading practices than those teachers who had much experience. The veteran teachers tend to be more lenient and focus on the whole child.

As an overwhelming similarity from all of the participants, teachers base their grading practices on the school’s expectations. Each participant shared that the schools have practices in terms of the weighted averages used to determine a student’s grade. For example, the most common breakdown was 40% tests, 30% quizzes, 20% classwork, and 10% homework. Teachers believed they had the most flexibility within the classwork component; however, Tyler reported, “Within those categories, I try to make sure that the things I value most are the ones that are going to be the highest percentage.”

**Influences**

Teachers are influenced on a regular basis in all aspects of teaching. Usually these influences are initiated from their student-teaching experiences, working with mentors and coaches, and collaborating with other teachers within the same department. In this study, some of the participants have expressed that experience has been the biggest influence. For example, Tyler reported, “The very first year I taught, I didn’t have a lot of grades because I felt that they were learning and that was all that mattered. I quickly realized that they needed to have more, so one grade wouldn’t hurt them.” Rachel too, credits experience for how her grading practices have evolved. She stated,

The longer I’ve taught, I learn as I go. It’s something that has evolved. I didn’t take participation and effort into consideration at the beginning of my teaching
experience, but now, I look at what they’ve put on paper and what I’ve observed in class.

Several teachers reported that their first decisions about grading practices echoed the practices of their teacher-mentors when they first began teaching. Michelle reported,

During my first year, there was another language arts teacher that helped me a lot with instruction and grading assignments. She was the one that got me started on focusing on a specific genre every nine weeks, and I’ll either give a test or project. The independent student reading logs came from her too, and we all count them as a quiz grade. Other than that, there is one other teacher that I get ideas from still.

Cooper concurred that her mentor had a role in helping her develop her own grading practices, but as her teaching experience expanded, so did her philosophy on what influenced her in terms of grading student work. She reported,

When I first started, I just looked at my mentor’s practices. Then, I realized I don’t give as many quizzes as she, so I wanted to downplay the percentage. As I continued teaching, I felt like the students were not getting enough credit for the good things they had accomplished in class. This is what put the participation grade in class up to a higher percentage.

Therefore, Cooper’s influences have evolved over time based on her knowledge of her students. For example, Cooper capitalizes on what the students are doing well during class time. With some students, Cooper recognizes that completing homework is not a priority at home, and chooses not to penalize the students when home situations don’t support a learning environment. As she reported,
Students still get 50% credit if they don’t complete their homework. To be totally honest, there are certain students that I will assign a zero because they typically do their homework all the time, and it is unusual when they don’t. Those are the students that I know don’t have a reason. However, there are certain students that I know about their home lives, or I know what responsibilities they have outside of class, or even some of my special education students, so I will not assign them a zero. For the student who just didn’t feel like doing it, I will assign a zero. It isn’t equal, but in my experience, it’s fair. I think you have high standards for everybody, but I don’t think they can be the exact same standards for everybody. I feel my focus has shifted. I hold myself responsible for more things. It’s not that I hold the students less responsible, but I’ve gotten to know what they are capable of being responsible for, and what they are not.

Smith believes that her grading influences come from knowing what she wants the students to learn from the unit or assignment. She approaches the planning of instruction, assessment, and grading with an outcome in mind of what she expects from the students.

**Academic Factors**

Powerful discussions took place when teachers were asked to describe which factors they considered academic, meaning which factors contribute to student achievement. It was unanimous among the participants that student projects, tests, and quizzes are all considered academic factors. The factors which were most debated and which teachers found difficult to label were classwork and homework.

As Lee stated, “I consider academic factors to be those that are based on SOL standards – those that the students need to know in order to move on to the next level.”
They include assessments, reading comprehension, vocabulary and certain parts of the writing process.” Rachel clearly defined the academic factors by listing classwork, homework, quizzes, and tests. Although she recognizes that student attendance and behavior affects performance, she does not consider those academic. Tyler agreed with Rachel in terms of quizzes and tests, by stating, “Test grades are certainly academic, because it’s always based on what they understand and what they know,” but she does not include homework in her list of academic factors. In terms of classwork, Tyler believes that “Classwork has an element of achievement, but it is often effort-based, and therefore not graded.”

Cooper, on the other hand, believes that homework falls into both categories of academic and non-academic. She supported her belief by stating,

I think it’s academic in that it shows if you complete your homework because you’re getting practice and hopefully it will show in your tests. I think for that reason it is academic, because I never grade for accuracy on homework; it’s always just completion. But homework completion, I think falls under the non-academic factor too, because some students have homes where they don’t have to do homework or have a good place to do it.

Both Ann and Smith believe that academic factors should measure student content understanding and should be reported as achievement grades. Lee makes the distinction by stating, “To me, academic achievement also depends on the spectrum; where the student started, and where they are now. If a student began the school year reading at a fourth grade level and is now reading at a sixth grade level, that’s academic achievement.” In support of Lee’s statement, Smith concurs by stating, “I would hope
that we will get to the point where we look at where the student was and look how far this student has progressed.” Edna believes that factors that qualify as achievement should be based on student work. She stated,

You can’t grade a student down because he annoys you. I am very cognizant about student work. A student can hand in the roughest thing in the world, but if there’s effort in it, then you’re going to get as good a grade as the kid who is the gifted artist who can do something that is just unbelievable.

Edna continued with a story about her own child’s effort on a school project that was completed solely by him. He met all of the expectations of the assignment, but because his strengths did not lie in creativity and neatness, Edna believed that her son’s poor grade was a reflection of presentation and not a reflection of meeting the expectations of the assignment and demonstrating student effort.

“Ideally”, stated Tyler, “all grades should reflect academic achievement. However, in the world in which we currently live, I don’t think it’s fair and necessary because each student comes to the school with such a different situation.” Tyler’s response was similar to what Cooper reported earlier about a supportive home environment in terms of supporting student work but also supporting teacher expectations. Tyler concluded with,

I give students more opportunities to turn in late work. I hate taking late work, but the reality is if I take it, it saves me a lot of hassle. So, ideally, grades would reflect academic achievement, but in the real world, I don’t know that it is always the case. I don’t want to deal with the headache of having to argue with a parent over a late assignment.
Similar to Tyler’s thoughts, Mary stated, “I think that grades really should reflect academic achievement because sometimes we inadvertently pad grades. Grades should be reflective of student’s actual abilities. When a standardized test scores are different from classroom grades, there’s an obvious disparity and not fair to anyone.”

Michelle admitted to being at odds with defining which factors contribute to academic achievement, and which are considered non-academic. She stated,

I feel like language arts is so open-ended, and I think that sometimes when no one is sure where we should put something, it goes into language arts because everything can fit into the language arts curriculum. So, I think it’s really hard, at least in our subject, to only count academics because some of what I’m looking for is work ethic because it will be important later. I try not to be arbitrary about some of those things, but sometimes I am arbitrary.

Non-Academic Factors

The discussion about non-academic factors was very enlightening. Contrary to how research defines non-academic factors, the teachers describe some of them as academic factors since they contribute to student achievement. Such factors include student work habits, participation, behaviors, effort, responsibility, and organizational systems. One common theme that evolved was the use of non-academic factors as a cushion, especially when a student is on the borderline between two letter grades. Lee summarized this thought when she stated, “If I need a little extra cushion or something like that, then behavior, effort, and participation play a factor.” When Cooper has students on the borderline, she takes into consideration whether students took advantage of opportunities to earn bonus points. As she stated, “Every student was given the
opportunity. Whether or not they chose it is their responsibility regardless of their home life because most of them could have done what I was offering in class.”

Tyler emphatically disagrees with use of using organizational systems as part of a student’s grade. She recalled a graduate course in which the instructor helped her understand that student organization has no place in a student’s overall grade because although organizational systems may impact how a student learns, studies, and prepares for class, it is not related to the student’s overall achievement. Tyler also believes that grades on organizational systems “help the students that were already doing well, and it hurt the students who were already not.” In her opinion, “If organization is a problem, then they are already suffering those consequences in the assessments. I don’t need to penalize them further because they can’t get it together. I try to allow for students who learn in different ways to have those needs met.”

Student Behaviors

All of the participants agreed that student behaviors impact student learning, referencing this topic 11 times throughout the interviews. “Student behavior definitely impacts performance in school,” reported Rachel, “but I don’t see that as academic. Behaviors concern me because misbehavior leads to students missing work, so it impacts their achievement.” Edna concurs that student misbehavior can affect a student’s grade and broadens the meaning to include missing assignments or not listening. “If the misbehavior impedes the learning of others, then that is another situation, but that still does not affect the grade. I’ll just take care of that in another way,” continued Edna.

Smith also reported, “I tried really hard as a classroom teacher to not let behavior influence my grading. Just because you could be a really, really smart student and have
this quirky thing about talking out and talking over somebody else, that shouldn’t affect your grade. Your grade is your grade.” An interesting and different perspective about student behavior and teacher responsibility was given by Fred. She stated,

    I don’t think student behaviors should be reflected in student grades. I have a student who is an absolute terror behavior-wise, but he is very bright, and he performs very well. He typically gets A’s in almost everything that he does because he does care about his grade. Behaviorally, he’s a huge disruption and a big problem. I think behavior is something we teach; it’s not something that we are grading, but it is something that we teach, model, and expect. I don’t think it should be part of their grade, but it is something I feel a responsibility for teaching.

Mary’s response was similar to Fred’s with regard to teachers taking on the responsibility for student behaviors. She stated, “I don’t think that behaviors and discipline should enter into grades, and I know that’s a very slippery slope because we would love students to be on task. So, it’s a challenge to try to manage the behavior but have the academic performance be reflected.” Mary continued to say, “I think it falls into the planning category – you have to plan for keeping those students motivated, on task, and involved because that’s part of what we do as teachers.”

**Student Effort**

Effort was probably the most used word in all of the interviews with 22 separate comments coded throughout, noting minor differences about teacher’s use of student effort in terms of grading. This suggests a shared philosophy among the participants that student effort is valued as an important consideration when grading student work. The
participants used the words *effort* and *progress* interchangeably during their discussions. Once Lee gets to know her students, she looks to see if student effort is matching their ability. This gives her a bigger picture and better understanding of what they have learned, and she stores that information away in case a student is on the borderline between letter grades. Rachel views student effort as the amount of student thought, preparation, and work that went into completing an assignment. Recognizing that her method is very subjective, Rachel assesses student effort over time on the spectrum from bare minimum to above and beyond.

Edna agreed that effort is particularly important in validating middle school students’ work over time, to know that the students, not their parents, completed the assignment. Edna finds this to be more important in student writing pieces. Tyler finds that effort allows for wiggle room when assessing student work. She stated,

> For example, I have one student who is incredibly bright and he will do exactly what I ask for in the rubric – nothing more and nothing less. I might dock him a little because he’s done it, but he hasn’t done it with the excellence that he should. In this case, there’s a bit of wiggle room for effort.

Cooper has a similar philosophy about student work demonstrating effort when she stated, “Effort is a big factor for me. It’s progress over time with a little bit of enthusiasm or ambition with the assignment or project. I want them to try.” Mary is in alignment with Tyler and Cooper when she reported,

> Effort plays into my grading when it’s springtime and I have students rushing things. I can tell if a student didn’t take a lot of time and care; didn’t use the strategies that we embedded in our generalized instruction. I can detect that the
student didn’t give that piece of assessment the right seriousness. When that occurs, I continue to work with the student and his or her assessment, because I want the student to master the material before moving on.

When discussing the factor of effort in student writing pieces, Mary stated, “Effort would actually be graded in a writing assignment in the areas of elaboration and presentation. Careless errors, sloppy writing, and writing minimalistically end up going into the presentation category.”

Ann views effort as a student motivator when she reported, “Effort should be part of the grade. It should not be a major component of the grade, but if you don’t reward a student for trying, they don’t earn the great grades that other students earn. What’s going to motivate them to try?”

**Student Responsibility**

A much-debated topic, student responsibility was referenced 15 times throughout the interviews with opposing statements. It was a very clear finding that the participants were divided about including student responsibility as part of a student’s grade. Six responded strongly that student responsibility should be part of the grade, while the other four participants do not feel that responsibility should be reflected in students’ grades even though they recognize the importance of student responsibility.

The group of teachers in support of using responsibility as part of the grade made similar statements about student participation in their share of learning. Rachel’s comment, “I think students have a responsibility to do their share of learning; retaining material, and doing what is assigned. I think that responsibility speaks volumes when it comes to showing what you know,” was echoed by Ann’s statement, “Grades reflect
responsibility on the part of the student, and I think it’s important, but responsibility grades will not fail a student.”

Lee associates student responsibility closely to student effort when she said, “Student effort and responsibility are huge. I told my students before that I would rather have a student that works well but doesn’t always get that hundred than a lazy student who just works the minimum. Yes, I do take into consideration responsibility and effort.” Edna associates student responsibility with completing and turning assignments in on time. “It has come into play at some point. It’s not academic, but a student must show some willingness to put forth the effort to hand in assignments. If a person consistently doesn’t hand in anything, they may not pass.”

On the opposing side, Fred is passionate when she spoke about student responsibility and how it affects student performance.

Responsibility is an important functioning skill for a student to have, and for a worker in the world. So, I do think it’s something that we need, that we work on, and it is something that I talk to the students about until I’m blue in the face. But, I don’t think that it deserves a grade. I do keep after the students by using checklists, but it’s not a true measure of whether they are learning or not.

Fred readily admitted that she does not grade students on responsibility but recognized that other teachers do. After some reflection, Fred continued with the statement, I would like to find a way of not grading responsibility, but some better ways of keeping track of those kinds of skills that help you in the work world. I think if you totally divorce those skills from assessment, I think it really depends on what you are assessing. If I am assessing responsibility, then I am going to, but I don’t
want to give it a grade. It is that constant dichotomy. I try to teach them as
good people. I want to make them better people. I want them to be more successful in
life, and to do that, I need to talk some about responsibility. I need to talk to the
students about getting things done on time.

Michelle is also part of the opposing camp, yet she was conflicted when discussing
student responsibility in terms of grading because she feels at odds with her own
practices. “I don’t think that student responsibility should be part of their grade, but I
know that I include it. It doesn’t usually count for that much, but if they are handing in
things late, then I penalize them for it.”

Factors Deemed Most Important

Teachers were divided on which grading factor they considered the most
important as it related to student work and achievement. Fourteen references were coded
from nine participants. Six of the nine believe that student classwork that demonstrates
progress and development over time is the most important grading factor. This important
finding suggests a common and comprehensive set of beliefs that the use of formative-
type assessments are the most meaningful in terms of identifying students’ strengths and
challenges, assisting the teachers in curriculum planning and instruction.

All six teachers stated that through classwork assignments, they were able to
interact with the student work while it was in progress, providing immediate feedback to
the students on their improvement. For example, Edna stated, “If the students do a quick
little grammar assignment that is part of a writing assignment and I see that everyone is
getting this wrong over and over again, I may do a lesson just on run-on sentences and
fragments.” Edna elaborated on her assessment process during classwork assignments by
sharing that once lessons are taught on specific areas, the expectation is that students will no longer make similar mistakes. She stated,

On writing assignments, I take about 20+ minutes to grade a two-page paper, because I will read it for the first time for the content, and then I reread it for grammar. I may not take off for everything, but I mark everything. You can’t put a B on a paper and just say it’s a B paper and not have anything on it. Even with an A+ paper, I still make comments on it.

In Tyler’s response, she gave a quick answer to the question about which factor she placed the most importance. She stated,

Classwork assignments are the most valuable part of all of the assignments because it’s my chance to catch them before they make mistakes. And it’s not even when I’m grading, it’s when they are asking me questions. I think with homework, too often the students just dash it off on the bus on the way to school, so I don’t feel that it is necessarily the most helpful assessment.

Fred concurred with Tyler’s approach to using classwork as a method in assisting the students with their work when she stated,

I feel like anybody can study for a test; anybody can study for a quiz. Some are better at studying than others, and that is a factor in a quiz or a test. What I really look at in the classwork assignments and day-to-day writing is development. If I see development, then the grade is a good grade.

As a stand-alone response, Cooper places the most importance on the factor of effort and what that means in her classroom.
Effort can be reflected differently; in how well they do on a project; how much they turn in their homework; how well their homework is done; their participation in class. I think it is very broad and not very specific, but I think effort is what I place the most focus on. If I know a student is trying really hard, but just not getting it, that is different to me than a student who is getting D’s because they are just not trying at all.

The two remaining participants place the most importance on student test grades. Mary and Ann both agreed that student achievement is reflected in their students’ test scores. As Ann stated, “We want the children to master the content. Of course, we want them to know how to read and comprehend, that’s our major goal in the language arts curriculum.”

**Summary of Factors Considered when Grading**

A significant amount of time in each interview was devoted to the theme of grading factors. The researcher noted that participants spoke passionately about grading factors and grading practices, ranging from the influences in their professional careers that helped define how they currently grade to those factors they identify as most important. At some point it appeared to the researcher that the teachers became uncomfortable in discussing their choices in using academic and non-academic factors. The researcher’s field notes suggests that this discomfort may have been due to their necessity to defend their positions, which gave the researcher the impression that perhaps the teachers have had to defend their practices in the past.

When grading student work, teachers considered many factors, ranging from student behaviors to student performance on tests and projects. When the researcher
considered the discussions that resulted in identifying grading factors, it was apparent that one important factor emerged – student effort, indicating a shared viewpoint that student effort is a valuable component when grading. The earnest discussion about the importance of using non-academic factors, such as student behaviors and student responsibility led the researcher to believe that the participants have had practical experiences in which non-academic factors have resulted in both positive and negative experiences for students. Once grading factors were discussed at length, the flow of the discussions naturally led to how grading practices are used in formative and summative assessments.

**Assessment**

Although the words grading and assessment are often interchangeable, the two are distinctly different. Classroom assessment is used to gain an understanding of what the students know in order to make informed decisions about how to move forward. In this study, there were 49 comments from the ten interviews which reference classroom assessments. Teachers use many different strategies to accomplish this. The strategies can be broken into two main types, formative assessment and summative assessment.

**Formative Assessment**

In education, the term formative assessment has become widely used. More importantly, the use of formative assessment has increased in its understanding and practical use by teachers in the classroom. Although one teacher readily admitted that she was unsure of what the term formative assessment was, once it was explained to her, she understood that her practices in the classroom can be defined by this category. All ten participants spoke passionately about using formative assessments in the classroom.
In this study, a clear finding emerged about the use of formative assessments in the language arts classroom, suggesting unity about the importance of formative assessments. The participants believed that formative assessments were mostly used during the writing process in their language arts classrooms. They include warm-up exercises, independent or group classwork assignments, and on-going writing assignments.

Lee demonstrates a strong understanding of formative assessments when she stated, “I use a lot of formative assessments in my classroom because it’s that one topic that we are focusing on, especially when I’m teaching writing and the separate traits. I’m looking for that one thing, not a cumulative point.” However, Lee doesn’t only use formative assessments during the writing process; she uses it in reading lessons as well.

I use it during reading with the different elements of non-fiction and fiction – it is so ongoing. When we use literature circles, I bring formative type of assessments in at any time. During literature circle discussions, I really get to know my students better, and higher level thinking can come out.

Lee supports the use of formative assessments because she believes that students need comments and an explanation, not just a score. In support of Lee, Rachel stated, “Informal assessments are more powerful and more informative than formalized assessments. I make notations on the students’ work, and I get a better understanding of where the student is in their progress.”

Edna was passionate about her response in using formative assessments with her students.

I confer with the students every single time on major writing assignments which takes a very long time. I will sit down with them for at least three minutes to say
this was really good; you need to work on that. I confer with them because how many students will actually sit there and read the comments and figure it out? But they will listen to me. I tell the students that they are going to be awesome writers by the time they finish the year because I am so picky.

Edna’s attitude about students improving their work over time is demonstrated when she stated, “I want the students to re-do their work because just getting a bad grade is a license for them to continually do the same thing over and over again. Giving them a second chance gives them a reason to do it right.” She concluded the discussion about the use of formative assessment with,

Since formative assessment is to provide feedback for the students to improve, that is why you have to sit down and conference. When we do differentiated readings, and I want to find out what the students understand, sometimes they can’t articulate some of the things in writing, particularly in certain classes. But if they sit down and I’ll ask questions about the meaning, they’ll nail it if they can just talk about it. It’s amazing. I talk with the students a lot, and I get them to talk to me too – you have to, particularly in language arts. Students are starved to have adults talk to them. It’s huge to have somebody validate something that they have done, or pick out things. I have to tell them that they are good, and then they become good.

Tyler and Cooper both agreed that the only way student work will improve is through providing meaningful feedback through constructive formative assessments, not just depending on assigning grades. As Tyler stated, “Let’s be honest. There will never be a true cumulative assessment in writing. At eighth grade, I can’t expect them to be college
level writers, so there’s always going to be improvement. Giving them a grade isn’t really fair.” Cooper stated something similar when she commented,

I don’t grade for a grade on their first writing; that wouldn’t be fair. I collect rough drafts and I spend time making comments, pointing out what they need to focus on – true feedback that takes forever to write. That is huge, and it’s huge for the students who use it. The most frustrating part is when they haven’t changed a thing on their final copy.

Cooper also uses formative assessments in her reading instruction. She offers extra practice as classwork or homework assignments to help students improve their reading comprehension. “If I know a student is weak in something, I feel like they should be doing extra work more than others.”

On the other hand, Mary finds that providing comments on student reading work can be meaningful and motivational, but finds it difficult when the focus is on objective work, primarily multiple choice opportunities. “I find it hard to give some kind of comment to something that is just so black and white.” However, Mary’s belief changed when she spoke about student writing. “Comments do become very crucial when writers are drafting. Every little snippet and nugget is so treasured and valued as far as comments go.” Ann was in alignment with Mary’s statement when she said, “The students really look at the comments, not the numbers because the comments are more meaningful to them.”

Fred’s perspective about the use of formative assessments is common to the other participants. Not only does Fred gather information about student learning through quick
day-to-day assignments, but she emphasizes the importance of the reading and writing conferences that her school supports.

We have at least two reading conferences and at least two writing conferences with each student each marking period. I keep a notebook with running records, and I usually mention what they’re reading, what they like about it, and a comment that they may have made. They actually start to think more about it and explain why they are reading the book, and what they’ve learned from it, instead of giving them a test on the book when they finish reading it.

In terms of writing, Fred recognizes the benefits of building relationships with her students when she said,

With writing, I use talking points and ask questions about their ideas and characters. Students like to write narrative stories about themselves. Language arts teachers are at a definite advantage because we really get to know the student on a much deeper level than maybe a math teacher because we read what they write. I get insight into what they are like and how their family approaches problems because they write about those things. It’s a big responsibility, but it’s very rewarding. For me, it’s the best part of language arts. It makes all the papers and all the hours of reading things very worthwhile.

Overall, the participants share similar philosophies in recognizing that formative assessments are beneficial in providing feedback to students to help them make decisions about improving their learning. As Smith summarized, “This discussion about formative assessment is so timely because it has been an instructional goal for our entire school. We want to move beyond just looking at standardized scores.”
Summative Assessment

Summative assessments are the traditional type of testing known in education such as unit tests, chapter tests, and standardized tests, where the outcome of learning is a fixed level of accomplishment. Properly using a mixture of summative assessments and formative assessments is the key to helping teachers make instructional decisions. The use of summative assessments was not a highly-discussed topic with the participants, only citing ten comments throughout the interviews. All participants had similar beliefs about the use of summative assessments, suggesting a shared opinion. They all recognized that summative assessments are part of a school environment, especially when it comes to standardized testing, but all of the teachers agreed that relying solely on standardized test scores does not reflect a true picture of student understanding nor helps with making instructional decisions.

Rachel, apprehensive about summative assessments stated, “I think there is a need for summative assessments to hold students accountable for what has been taught; however, I don’t like the weight that is placed on one single score. There is too much pressure put on the students and teachers to perform on a one-day test.” Edna, who recognizes that much of the student performance in the language arts classroom is subjective, understands the importance of using summative evaluations in order to have objective tests, but she maintains that she adds an element of subjectivity, such as an essay, to look for critical thinking from the students.

Mary’s approach to helping students reach success on a summative assessment comes from years of fine-tuning her practices in the classroom.
We practice a lot. We try to incorporate all of those different elements. We try to read it and write it and speak it and share it. Then I turn them loose on their own. Hopefully, I have given them enough support and stability, which I eventually scaffold back from, to let them really show what they can do on a summative evaluation.

Smith’s reaction to the topic of summative assessments was very similar to Mary. She uses formative assessments throughout the year to prepare students for summative evaluations, and she agreed that summative evaluations are to be used to provide a final grade as an outcome of learning.

Conversely, Ann does not believe that summative evaluations should represent students’ final grades because she believes that some students need the open-ended option where they can explain what they have learned. “Students can do poorly on a summative assessment, but there are just portions of that summative assessment that they do poorly on. So, I still go back and look at those critical attributes that they missed.” Summative assessments are considered necessary where Ann teaches, yet she still continues to analyze the results of the summative assessments and make instructional decisions, much like the use of formative assessments. On the other hand, when the topic of summative assessments was discussed, Fred simply replied, “I don’t use summative assessments; I use a combination.” She explained this response more thoroughly when she elaborated with,

I think within the nature of education, you have to use some summative assessments, and you have to say at this point this is how the students performed.
Depending on circumstances, some students may have not done well, so I may give it to them again next week. That’s part of what we do.

In addition to using traditional unit tests, Tyler and Michelle identify with using projects as summative assessments, alleviating the concern that student performance on a one-day test provides a final grade as an outcome of learning. Each of them described a similar approach of using tiered student work that would eventually be used to create a final product, such as a unit project. All of the tiered work was described as formative assessments – mostly student work that was completed within class in order for the teachers to provide feedback before the work became part of the project.

Cooper summarized her belief, as well as what the other teacher-participants believed about using a combination of assessments to reflect student grades when she stated,

I agree that we need to use some summative assessments to provide us, the teachers, a grade of student learning. I think they are necessary to look at as straight up academic achievement – the students can do it or they can not. So, I think they are necessary, but I don’t think that should be the only thing factored into their final grade.

**Summary of Assessment**

Assessment is often interchangeable with grading, which was evident when the participants spoke about student assessment. Most participants recognized that daily assessments occur in the form of questioning, warm-up exercises, and homework assignments, all known as formative assessments. However, during the interviews, the
researcher’s field notes suggest that the ideas of formative and summative assessments were not clearly understood by all participants.

Very little was discussed about summative assessments, and most participants identified summative assessments with standardized testing; however, the greater discussion took place regarding the use of formative assessments. For some participants, they readily admitted that the term formative assessment was relatively new to them, and that they were still learning how to properly use formative assessments as instructional tools. This discussion was the ideal segue that led to meaningful and practical professional development opportunities.

**Professional Development**

In every school division, financial and human resources are set aside for professional development of all staff – teachers, paraprofessionals, support employees, and administrators. Professional development allows all stakeholders to grow in their area of expertise and remain informed about best practices. However, professional development in education does not begin once teachers are hired to teach, it begins in teacher training courses when they are college students and continues throughout their careers as classroom teachers and leaders.

**Teacher Training**

Only 13 references were made throughout the interviews concerning teacher training as it relates to grading practices. It was noted that only one of the ten participants received training in their undergraduate courses in how to grade student work. All of the teachers remember having some type of training in how to develop assessments, but Mary was the only teacher who remembered having lessons in teacher
preparation courses about how to grade assessments once they had been developed. Mary said, “We had a very hands-on approach to making sure that we validated whatever we chose as a component in the grade book. We were introduced to the concept of having grades weighted by percentage, which was a big focus.”

Smith reported that she learned strategies in classroom management, behavior management, and lesson planning during her teacher training while in undergraduate school, stating, “Training in technology, creating assessments, and determining best grading practices was during my first year of teaching.” Tyler has a similar viewpoint when she stated,

Beginning teachers are on such a learning curve. There are so many things to learn. I feel as if students who have a first year teacher should almost get a second chance at the year, because there’s so much the teacher is learning. I think it would be very helpful to have training in areas such as grading practices – if someone really sits down with a new teacher and explains how it works.

All of the teachers believe that during their student-teaching experience as well as professional development opportunities within their schools and school division, they have learned valuable lessons about grading practices as it relates to current measurement theory and recommendations.

Moreover, teachers credit the trial-and-error approach in determining what works best for them and their students in the classroom. Edna reported, “Back in the 1970’s, they gave us some recommendations, but throughout the years of just trial and error and asking other teachers for their advice, is the way I have determined what works best for me and my students.” Ann concurred by stating,
I wasn’t prepared at all on how to establish my own grading practices, but I learned on the job by trial and error, and discussing topics with other teachers. Over the years, my practices have changed. When I first began teaching, I focused on the number grade. I don’t know if that was inexperience, or just not knowing my students, or developing relationships with my students. I learned over time that my grading practices were very black and white. You either do this and you get an A on the test, that’s it. This has changed. I think that if you look at the whole student, you can see where their strengths and weaknesses are, and focusing on that improvement is really achievement.

Fred reported that she was not trained in grading practices during her undergraduate work, but through participation in her master’s work and through a funded writing project, she learned much about assessments and how to grade them. She reported, “It was eye-opening. That was the first time I had read a lot of research on scaffolding, frameworks, and formative assessment, and it was like wow, where has this been all my life? Most classes don’t focus on developing assessments or how to grade them.”

When the use of rubrics became a popular way to determine a student’s grade, teachers felt that it alleviated some subjectivity on their part. As Smith reported, My first exposure to any kind of performance assessment and how to grade that came during my first year of teaching when I was introduced to rubrics during a professional development training session. I see a real need for using rubrics, and it’s an easier way to help justify student grades, especially with parents. There leaves little room for subjectivity because a rubric is based on objectives and criteria.
Teacher Conflict with own Grading Practices

Teacher reflection is important. Throughout each school day, teachers make decisions about planning, instruction, assessment, and classroom management. Some decisions are well thought-out, and some demand instant resolution. Reflection is used by teachers to make improvements, and in the case of grading practices, teachers find that there is room for improvement. Citing 17 different comments in this sub-theme, the participants could recall at least one time in which their practices were in conflict with their own beliefs or the policies in place at their schools.

Rachel recalled a time when she was at odds with her own expectations of grading when a student performed poorly on a unit test, but his participation and answers in class preceding the test demonstrated understanding of the unit. As she recalled, “Do I mark it wrong, because it’s not the answer I expected, or do I give him credit because I know that he knew the material since he could talk about it in class?” Rachel remembered that she did indeed give the student credit because she stated,

My practices are going to vary from activity to activity, from student to student, from year to year. I learned from that experience that I have to revisit the way I do things; that is especially true with grading. There will never be growth if everything stays the same; there have to be modifications.

Similar to Rachel, Michelle has conflict when student performance does not match capability and intelligence. She believes that student work habits, such as turning assignments in on time and following directions, negatively impact an overall grade because a student may have the intelligence but not the work habits to produce quality work.
When Edna has conflict with her own grading practices, she has learned that practices and policies need to be flexible because she remembers that she is teaching young people who sometimes are going through quite a bit at home. Edna credits her compassionate side of working with children due to a memory from when a professor was compassionate with her during her college years.

Tyler shared a similar story when she recalled using flexibility and compassion with a student who was out for two weeks due to a family emergency. She was able to excuse him from making-up a test because she knew that he was overwhelmed with work in all of his classes, and he had demonstrated, prior to his absence, that he understood the work.

However, Tyler is at odds with her own practices because she feels that some of the school’s grading practices do not match her own philosophies. For example, at Tyler’s school, students earn a minimum grade of 50 on quizzes, tests, and major assignments. Tyler reported,

I give away points with our current policy for students who I feel would not be as successful, so I don’t feel that the grades are an accurate reflection of what they understand. In some ways, I feel like I am rewarding students who just blow it off, because at this point in the game, students are smart enough to figure out this point thing. I don’t think the practice is fair. I feel like students who work really hard get the same grade as a student who didn’t work really hard simply because that’s the practice. So, I don’t feel like my grades are reflective of academic knowledge, but more a reflection of effort.
Conversely, Cooper recalled a few instances when she was at odds with her own grading because of her own subjectivity, mood, and time allotted to grade assignments. “I was way too hard on this one student when grading a project because it was midnight or later, and I was rushed to complete it. Grading is just so subjective to your mood and time. “In this instance, the student asked me about her grade, and after I reviewed it, I knew that I was too hard on her, so I changed it.”

The researcher thought Mary’s response was filled with humor when she stated, “I have conflict with my grading practices every single time because we’re not supposed to fail our students on homework, or homework is not supposed to have a terribly negative impact, but I want it to matter to them.” Since Mary feels that her hands are tied with some practices, she found a system that works for her. “I like volume because it gives an average, and an average lets me know if students consistently perform.”

Lee, too, was able to recall situations in which she was at odds with her own practices, but her response summarizes this section nicely on a larger scale. Lee’s concern is when students move from grade level to grade level, that their performance may be questioned due to the different standards and practices that teachers have in place.

It’s hard because you hear that a student did so well in the prior grade and made straight A’s, but the student is performing so differently for you. So, you wonder, how did the other teachers grade? I think learning about grading and assessment is good, because you don’t want a student to fail because of a previous teacher and they didn’t get all they needed, but they’re doing just fine. Teachers’ standards are different at every single level and with every single teacher.
Need for Professional Development

As described earlier, the teachers reported that since they were not trained in teacher preparation courses in the area of grading practices, they believe that a need exists in this area. Although one-day professional development sessions are currently serving this purpose, the teachers believe that more information about grading practices is imperative to assist with understanding the differences of academic and non-academic factors and how these affect students’ grades as it relates to student understanding and mastery of knowledge.

The topic of professional development naturally emerged once the participants discussed their lack of training in grading practices. Ten references were made by the participants, indicating an important finding. This suggests a common set of beliefs that as practitioners in their field of expertise, a great need exists for professional development in the area of grading practices. Most of the participants recalled instances when they adopted their own grading practices from trial and error. For example, Rachel spoke about an instance where she used a new assessment in class during the school year.

I used something new this year, and I wasn’t sure how I was going to grade it. I developed a rubric, and I had to play around with it – it really made me stop and think about what I expected the end result to be, what’s acceptable, and what’s not. So, I definitely think a course in grading practices would be valuable.

Cooper, who also did not receive training in grading practices while preparing to teach, credits her own experiences and time with colleagues in developing practices that best match her instruction, assessment, and how to grade the assessments.
I was never taught how to grade assessments, and I didn’t even know about formative and summative assessments. I don’t remember having a class learning how to best assess how students are truly performing. I’ve learned so much more about that just being here and through professional development. Fred’s experiences with student teachers have helped her form an opinion about effective teacher preparation programs.

College programs that send students on observations and practicums for all three or four years are more effective than those where student teaching is their first experience in the classroom. At this point, it’s too late for the students to learn about developing and grading assessments because they are expected to function and be competent. Many of the young teachers don’t know how to balance the types of assessments because a lot of what they know is that you take a test and you get a grade.

Mary was passionate about her opinion on this topic. “I think grading practices are just as important as any other factor.” She believes that since grades are the primary source of communication with parents that her grading practices validate her integrity in how she grades as a classroom teacher.

Lee concurred with all of the other teachers when she stated, “I absolutely think it would be great to have a course about what it means to grade and talk about key points in grading. To research and have a discussion about what is grading, how it reflects a child’s performance or growth, and how we measure that into what we need to know for state testing.”
Summary of Professional Development

Moments of clarity occurred during the discussions on professional development because it became rapidly obvious that teachers lack training in the area of grading practices. The researcher noted that the participants appeared not to have given thought to the fact that they had very little training in the area of grading practices until it was discussed during these interviews.

Professional development is an area that is rich in providing opportunities for teachers to learn new strategies, new approaches, and overall review curriculum standards and assessments. It is also an opportunity for schools to enrich their staff with new opportunities based on educational research in practical learning experiences to enhance instruction and assessment. When the researcher considered the discussions from the participants that resulted in identifying the need for professional development, it was evident that the participants completely understood their lack of training in the areas of grading practices and how it relates to assessments.

Summary

The emerging themes and subthemes help draw conclusions as listed in the next chapter. As evidenced by all of the participants’ comments and reflective thoughts, the subject of grading practices is thought-provoking and controversial. One common theme that emerged from the interviews is that having standard practices does not necessarily work when teaching students who are unique and individual. As Ann stated, “Every child is different. Every child learns differently. I have my set of goals. I have my set of expectations, but they are not always black and white. Life isn’t black and white, and I think you have to assess each child differently.”
Chapter V
Discussion, Conclusions, and Recommendations

This qualitative study produced a detailed description of how middle school language arts teachers consider academic and non-academic factors in their grading practices. The purpose of this chapter is to present a summary of findings, conclusions, and recommendations based on the data collected. This chapter is organized by the four major themes from chapter four: (1) purpose of grading; (2) factors considered when grading; (3) assessment; and (4) professional development. Recommendations for practice and recommendations for research conclude the chapter.

In conducting this study, the researcher was interested in examining how middle school language arts teachers consider the use of academic and non-academic factors when grading. The theoretical framework that guided this study was based on the composite model gleaned from the literature. Based on the field notes, the researcher believes that teachers spend more time thinking about instruction than they do on their grading practices. What appeared evident to the researcher was that teachers generally felt comfortable discussing their curriculum content, yet a discussion about grading practices appeared to be a bit uncomfortable and foreign since most teachers have not been asked to reflect nor discuss their grading practices. Once the researcher assured the
participants that the discussion was confidential and would be used to compare and contrast results, their comfort level appeared to increase.

Of notable interest, the findings of this study indicate a cyclical pattern on the subject of how teachers view their grading practices, which is in alignment with the literature. When the topic of grading practices was first introduced, the participants began their discussion with why they grade – identifying the purposes of assigning student grades. From there, the discussions naturally led to how they grade – which factors, both academic and non-academic they used in determining a student’s grade. When probed about what type of assignments students produce for their grades, the discussion turned to assessments – both formative and summative types of assessments. Once teacher-participants were asked how they determined their grading practices, this logically led to the topic of professional development in terms of teacher training.

**Purpose of Grading**

**Discussion of Findings**

During the discussion on the purposes of grades, the researcher noted that teachers do not spend much time reflecting on this topic. Grading is an expected responsibility; however, it has not warranted much reflection or discussion. Once the participants spent some time in reflection, they appeared at ease with the flow of the discussion, and their thoughts and opinions appeared to be truthful and reflective responses. The participants first began by defining the purpose of grades, which resulted in important findings. After comparing and contrasting the results, the researcher was not surprised that only four notable findings were defined as the purpose of grades: (1) to communicate with students and parents; (2) to provide feedback to students; (3) for
students to demonstrate progress and mastery; and (4) to adhere to grading policies and procedures. The first three purposes are identified by the participants as equally important, indicating dominant purposes, while the fourth purpose did not appear to be as critical.

The participants effortlessly pointed out that grades are used as a communication tool – to students, parents, and school administrators, suggesting a common set of beliefs by all participants. They agreed that communicating what grades represent at the beginning of the school year, along with reminders throughout the year on major projects and tests, alleviates confusion for the students and parents. Most of the participants agreed that using an electronic grade book assists with timely communication following graded assignments. Along with providing a number grade, the participants all agreed that another equally important purpose of grading is to provide feedback, again, suggesting a comprehensive opinion from the participants.

In this study, the participants found that communicating grades and providing feedback to students on their work provided a systematic approach in helping students improve their work, which ultimately enhanced student understanding. Their systematic approaches include communicating grade expectations through a course syllabus, communicating results of graded assignments to both students and parents in a timely fashion using electronic grade books, and providing meaningful feedback on assignments in order for students to improve their work. Meaningful feedback can be defined as written comments, suggestions, opinions, and views about ongoing student work. Relative to both important purposes of grades – to communicate and to provide feedback – the participants’ thoughts are widely supported in the literature.
In the broader literature, the purpose of grading is defined as a way to report results using communication methods, and to contribute to the learning by providing feedback to motivate students (Brookhart, 2009). Collectively, research conducted by Brookhart, Guskey (2004), McMillan (2008), Cross and Frary (1999), Austin and McCann (1992) and O’Connor (2007), indicated that providing communication to students concerning their grades assists with their overall performance because students benefit from accurate, specific, and timely feedback, giving them opportunities to improve their performance.

The interpretations of this study’s findings are consistent with the research conducted by Austin and McCann (1992), McMillan (2001), and Cross and Frary (1999) as it relates to defining the purposes of grades. McMillan’s study was comprised of a sample of secondary teachers, grades 6 through 12 in the subjects of English, mathematics, science, and social studies. Austin and McCann’s study consisted of a sample of only high school teachers in English and mathematics. Consistent with McMillan’s and Austin and McCann’s findings, this would suggest an overall finding that teachers in secondary schools, both middle and high school, identify that two main purposes of grading include communication and providing feedback on assignments.

However, in Holmes and Smith’s 2003 study about student perspectives of their teachers’ grading practices, students expressed concern that their teachers did not provide sufficient feedback on graded assignments. Although they agreed that grades were communicated to them, meaningful feedback was deficient, resulting in missed opportunities to improve their work. Furthermore, O’Connor (2007) challenges the belief that all communication is beneficial because he believes that grades misrepresent
student’s level of achievement and mastery due to poorly designed grading practices, resulting in ineffective communication. The findings of this study are inconclusive with O’Connor’s belief, largely due to the fact that the participants did not make a connection between how poorly designed grading practices could lead to ineffective communication, misrepresenting a student’s performance.

On the other hand, the findings of this study support the opportunity for students to demonstrate mastery and progress as another primary purpose of grading. The middle school participants found that student grades are a reflection of the student’s understanding and mastery of certain concepts. This information proves valuable to both the students and the teachers which ultimately assist the students with understanding their base knowledge, and assists the teachers with making instructional decisions. This interpretation is consistent with the literature by Brookhart (2011), McMillan (2008), Guskey (2004), Wormeli (2006), and Winger (2005). However, Winger cautions that students may not find the grades as valuable since they see schoolwork as a game they play for grades – stuffing themselves with information only to regurgitate it for the test, conflicting with the idea that students demonstrated deep understanding and growth. Sadler (2010) also believes that grades should reflect a reliable representation of a student’s level of achievement, yet questions whether all grades accomplish this purpose.

As the fourth identified reason for grading, this study indicates that adhering to grading policies and procedures is a factor to ensure consistency, yet this purpose was not as prevalent as the other three purposes. Consistent with literature, (Austin and McCann, 1992, Carifo and Carey, 2009, and Guskey 2006), developing common and usable grading policies are critical in making grading decisions fair, equitable, and effective.
In this study, the participants recognized that their individual schools developed common practices in terms of grading, yet they prefer to remain flexible in working with individual students. Therefore, they generally use common practices as guidelines, not firm and fast rules. The participants agreed that adhering to grading policies are sometimes difficult because they are ambiguous and vague, yet found that it worked to their advantage in order to remain flexible. Many of the participants discussed how the flexibility ranges from student to student with respect to their personal and academic backgrounds, and overall understanding of how the students learn. Clearly, participants in this study readily admitted that they do not adhere to division or school policies and common practices, and made grading decisions based on individual student factors. The interpretation of this finding supports O’Connor’s (2007) belief that grades are sometimes a misrepresentation of student performance.

Although the participants only identified four purposes of grading, the literature suggests that other factors contribute to the purpose of grading. For example, Guskey (2004) asked teachers to define the purpose of grades, and six main categories emerged. Of the six main categories, two materialized from this study – to communicate and to provide feedback to students. The other four that are listed in Guskey’s work include: (1) to identify students for specific educational programs; (2) to provide incentives for students to learn; (3) to evaluate the effectiveness of instructional programs; and (4) to provide evidence of a student’s lack of effort to accept responsibility for inappropriate behavior. Therefore, in this study, Guskey’s list of the purposes of grades is consistent with two findings, yet inconsistent with his other four categories, suggesting that communication and providing feedback are important purposes of grading.
Conclusions

As it relates to the purpose of grading, this study found that teachers of language arts in the middle school setting were similar in their beliefs about why they grade and the importance of providing feedback as it relates to literature and other research studies. An important conclusion was determined by the teacher’s discussions of the purposes of grading that is supported by literature and empirical studies. In order for grades to be effective, teachers should communicate results to students and parents in a timely fashion, while providing meaningful feedback on how the grade was determined. In addition, grading practices should follow grading policies, which sometimes have been identified as a flawed system.

Factors Considered When Grading

Discussion of Findings

During the discussion of grading factors, the participants appeared to defend their grading practices, as if they were being judged. This could be due to the fact that the teachers readily admitted to using flexibility in their grading practices which did not always align with the school policies that had been developed. Overall, the researcher noted that the attitudes by the participants appeared independent in thought, even though they could be greatly influenced by the current era of accountability and high-stakes testing.

One of the most noteworthy interpretations revealed from this discussion was the misalignment of participants’ grading practices and the recommendations developed by measurement theory experts. Measurement theory experts, as defined in Stiggins, Frisbie, and Griswold’s 1989 case study, revealed that of 19 measurement
recommendations, the sample of 15 secondary teachers were consistent with only eight of the recommendations and were in discrepancy with the other 11 recommendations. The findings of this study are consistent with Stiggins, Frisbie, and Griswold’s findings because the participants of this study were also in discrepancy with three of the recommendations listed in Stiggins, Frisbie, and Griswold’s 1989 study. The three recommendations in discrepancy are: (1) achievement should be the only factor in determining grades; (2) effort should not be used as a component of grading; and (3) formative type of assessments should not be used in grading.

After comparing and contrasting the results, four important subthemes about grading practices emerged: (1) influences on how teachers develop their own grading practices; (2) academic factors; (3) non-academic factors; and (4) factors that teachers consider to be the most important. The researcher noted that the lengthiest portion of each interview was on the topic of grading factors. The findings in this theme prove to be important, some consistent with literature, while others are not.

During the discussion on influences, the researcher noted that the participants were reflective and honest about whom and what they credit as their greatest influence in terms of developing grading practices. Two major influences materialized, both appearing of equal importance from the participants – their teacher-mentors, and their teaching experience. These findings are inconsistent with Guskey (2004), who contends that teachers develop grading practices based on what they are familiar, indicating that their practices as teachers are similar to what they remember as students.

The discussion of the academic and non-academic factors teachers use when grading proved to be the most referenced and most debated topic during all of the
interviews. This is consistent with McMillan (2008) and Dockery (1995) who agree that grading is a challenge for most teachers due to the variability in grading practices. In Brookhart’s 1993 study of 84 classroom teachers, she concluded that an assortment of variables were included in how teachers develop their grading practices, consistent with the results of this study. A closer look at both the academic factors and the non-academic factors assist with the understanding of the findings.

In this study, all participants agreed that academic factors included quizzes, tests, and projects; however, it was the discussion of homework and classwork where teachers differed in opinion. The researcher noted that the participants used the words *academic* and *achievement* synonymously when describing factors that reflect student understanding and mastery of material. This is why the participants had difficulty in labeling homework and classwork, because they question whether these two pieces of student work reflect student mastery of material.

Unequivocally, all teachers agreed that homework demonstrates the amount of student understanding of content, using it as a tool to make instructional decisions, but homework is often viewed as practice, not as demonstration of mastery. In addition, a recurring theme surfaced as it related to individual students. Many of the participants readily admitted to using different grading practices with different students based on their home environments. When the teachers suspected that homework completion, studying, and education were not valued at home, this influenced the teachers’ grading practices of student accountability.

Consistent with the literature, all participants use academic factors as a large portion of their grading practice. Collectively, research conducted by McMillan (2001),
Brookhart (1993), Cross and Frary (1999), and Stiggins, Frisbie and Griswold (1989) indicated that student grades should be a representation of student mastery of the content, which includes grades on assessments. Adding controversy to an already much-debated topic, Dyrness and Dyrness (2008) and Wormeli (2006) agree that only academic factors should be the exclusive components to account for student achievement. However, inconsistent with the literature referenced above, are the factors that are included in student grades that are not clearly defined as academic factors. The controversial factors are those that literature defines as non-academic: student behaviors, student effort, and student responsibility.

The interpretations of this study’s findings are consistent with the works of Guskey (2011), Brookhart (1993), McMillan (2001), Duncan and Noonan (2007), McMillan and Nash (2000), and Bursuck et al. (1996) as it relates to the use of non-academic factors when teachers develop their grading practices. In all of the above-referenced studies, results indicated that non-academic factors such as effort and class participation were used widely in determining student grades. Dyrness and Dyrness (2008) caution the use of using effort and other behavior factors because they mistakenly report a grade that reflects academic merit, but is muddied by behavior incentives.

In this study, several important findings emerged concerning the use of non-academic factors. Not surprisingly, all participants agreed that student behaviors affect student learning and the classroom environment; additionally, the participants agreed that student misbehaviors should not be reflected in student grades. However, the participants agreed that they use some non-academic factors, such as effort and participation, as a cushion when student grades are on the borderline. The interpretation of this finding is
consistent with McMillan’s 2001 study of secondary teachers, where the results of his study identified four distinct components of grading factors, one labeled as academic enablers, such as effort and participation, and a second component labeled as use of extra credit and borderline cases. Consistent with this study, participants’ use of effort and participation as a cushion falls into both of McMillan’s categories, academic enablers and borderline cases. This would suggest an overall finding that teachers in both middle and high schools agree that using non-academic factors such as effort and participation in the case of borderline grades are widely used and accepted practice.

As indicated above, a shared philosophy from this study indicated that effort is valued as an important consideration when grading student work. Teachers generally agreed that effort assisted with determining student progress over time as well as assisting teachers in making decisions about borderline grades. In fact, the word effort was the most used word in all of the interviews, and 22 separate comments were made by the participants with respect to student effort. The analyses of these results are commonly supported within the literature. For example, the interpretation of the importance of effort when grading student work is consistent with McMillan and Nash’s 2000 study of high school teachers and Randall and Engelhard’s 2009 study of elementary and middle school teachers, whose results support that student effort is important in developing grading practices. This would suggest an overall finding that teachers value student effort as a critical factor when grading.

The final non-academic factor that was heavily mentioned throughout the interviews was student responsibility, as it relates to students’ completing and submitting assignments on time. The findings in this area were inconclusive because the teachers
were divided about including student responsibility as part of a final grade. Those participants who were against using responsibility as part of student grade are supported in literature by Wormeli (2006). Wormeli is an opponent of using non-academic factors, such as student responsibility, in determining a grade. He cautions teachers in using subjective components because he believes it sends a message that student failure in responsibility affects student achievement, which misleads students about their level of content mastery. As mentioned previously, the teachers are influenced about students’ support or lack thereof in their home situations as it relates to making students accountable for submitting their work on time or at all.

The disparity in these results was not surprising to the researcher, for the topic of student responsibility is often debated, mainly due to the reasons that the teachers expressed. The researcher noted that the teachers are divided in holding the students responsible for submitting work, when they know that a lack of support at home may prevent the students from completing work. Therefore, some of the participants talked about maximizing students’ time at school to complete work, making them accountable for their work while supervised by the teacher. The participants agreed that this is a successful strategy to help students who lack support at home to be responsible and successful at school.

Although the participants in this study only identified three factors that they would categorize as non-academic (student behaviors, student effort, and student responsibility), literature supports that other factors may also be components of non-academic factors. In Bursuck’s et al. (1996) study, notebook completion, attendance and student organization were also factors in the category of non-academic factors, when
deciding which factors special education teachers used in determining grades. In addition, McMillan’s (2001) study of secondary teachers found that teachers also used student ability as a non-academic factor when determining grades.

The results of this study, with respect to grading factors, reveal that gaps exist between teachers’ grading practices and the recommendations made by measurement experts. Measurement experts outlined grading recommendations in Stiggins, Frisbie, and Griswold’s 1989 study. The 1989 study recommended that the only grading factors that should report student performance are those that relate to student achievement. As indicated in this study, the participants use an assortment of grading factors, ranging from achievement grades on tests and quizzes to non-achievement grades such as student participation, student effort, and student responsibility.

The results were divided when the participants were asked to determine one grading factor they consider most important. Two participants explained that test grades were the most important factor because it supports the curriculum and standardized testing within their schools. They believed that the test grades were a true reflection of student understanding and mastery. The two teachers who identified test grades as the most important appeared to be influenced by the rigor and demands of high-stakes testing that takes place for their students every year and did not give independent thought to their responses. Both are veteran teachers who realize that data analyses occur during the summer months by school administrators to determine student growth and overall achievement scores with respect to teacher performance.

Six of the participants determined that classwork was the most important factor in determining student grades, because it allows the teachers to give immediate and
meaningful feedback to the students on how to improve their work and ideas – this was especially true when teachers spoke about short-term and lengthy writing assignments. Only one participant identified effort to be the most important factor because she wants students to work to their ability and not settle for less than what they are capable.

Conclusions

As indicated by the qualitative data analysis, important conclusions were determined by the teachers’ discussions of the grading factors they use when developing their grading practices. Initially, teachers use influences from their teacher-mentors and their own teaching experience to develop grading practices. A mixture of variables, ranging from academic factors to non-academic factors, are used in determining student grades, identifying academic factors as those that relate to student achievement, and non-academic factors as those that relate to student work habits, such as effort and participation. The interpretations are consistent with the literature, supporting that teachers use mixed methods in determining student grades, and they lack consistency across schools and school divisions.

Assessment

Discussion of Findings

For many educators, assessment is synonymous with grading. Yet, in this study, the participants were able to delineate between the two types of assessment that are clearly defined in literature: formative assessment and summative assessment. The researcher noted that during the discussion of assessment, teachers freely admitted that they are eager to learn more about how formative and summative assessments work to the advantage of improving instruction and overall student achievement. The thoughts and
opinions shared by the participants appeared to be honest and forthright, yet still uncertain if their understanding of the formative and summative assessments were correct. Their attitudes about the topic were tentative because many commented that during recent professional development sessions at their individual schools, the topics of formative and summative assessments have been explored for greater understanding.

After comparing and contrasting the results of this study, two important findings emerged with respect to the use of formative assessments. All teachers agreed that the use of formative assessments is prevalent in the language arts classroom, especially during student writing exercises, which assists students in making decisions about how to improve their work, and assists teachers in making instructional decisions about curriculum pacing.

The teachers agreed that they use a systematic approach in providing immediate and meaningful feedback to students when they write, which helps the students improve their work. The exercises usually begin with generating ideas and developing an outline. At this point, teachers begin the process of providing comments and meaningful feedback to help the students with the direction of their writing. Once the rough drafts are written, teachers explained that they work quickly to provide meaningful feedback about the students’ initial writing, offering ideas on how to improve their work, focusing on the traits of writing – voice, style, connections, and mechanics. It appeared that all of the teacher-participants took pride in discussing how they use formative evaluation to shape the student’s work for the final product stage, since they all believed that students benefit more from comments than merely an assigned grade.
When discussing how formative assessments assisted the teachers with their own curriculum pacing, their ideas moved beyond the writing component of language arts. The teachers expressed that the students’ responses to classroom discussions and questions, and their answers on homework and classwork exercises provide considerable value in determining if students understand the material before they were asked to demonstrate mastery on a quiz, test, or project. Furthermore, the teachers believed that one of the greatest benefits in using formative assessments for student work is the method in documenting individual progress.

The interpretations of this study’s findings as it relates to formative assessment are consistent with the literature. Popham (2008), Brookhart (2009), and McMillan (2008) all define the purpose of formative assessment as the process of collecting evidence to help students make informed decisions about how to improve their learning. In addition, Sadler (1989) and Ames and Archer (1988) furthers the definition of formative assessment by explaining that teacher feedback assists students in identifying the gaps between their current level of understanding as compared to the criteria that meets expected performance and that the emphasis should be placed on continuous improvement, not necessarily the final outcome.

Marzano and Heflebower (2011), Wormeli (2006) and McMillan (2008) describe that formative assessment should be effectively used as a cyclical teaching tool. Wormeli emphasized that just as students are held accountable for their work, teachers are held accountable to the students. The results of this study, with respect to using formative assessments to inform instruction, reveal that gaps exist between teachers’ practices and the recommendations made by measurement experts. Measurement experts believe that it
is the teachers’ responsibility to analyze student work for improvement and make instructional decisions that do not advance the students until they demonstrate understanding and achievement. As indicated in this study, the participants’ descriptions of how they use formative data were not conclusive as it relates to reteaching.

With respect to summative assessments, most of the participants in this study shared similar opinions, noting important findings. Participants supported the use of summative assessments to evaluate student mastery of content, noting that the fixed level of accomplishment demonstrates levels of student understanding. The teachers agreed that summative assessments are a traditional mode of assessing student work at the end of teaching units; however, they rely on formative assessments prior to summative tests to properly prepare the students. In addition, some of the participants found that using summative assessments is a method in holding the students accountable for their work.

Another widely-shared opinion from most of the participants was the process in which teachers cautiously used scores from summative assessments to make instructional decisions, especially when the summative assessments were standardized tests. The teachers all recognized that standardized test scores are only one component of student understanding, and in order to make sound instructional directions, they heavily relied on additional evaluations beyond standardized test scores. The participants’ thoughts are widely supported in the literature, providing evidence that the interpretations of this study’s findings, with respect to summative assessments, are consistent with the literature of Brookhart (2009).

Literature defines summative assessment as evaluations that measure student work over time to gather evidence in reporting individual student grades (Brookhart
2009). In the current era of high-stakes testing, Brookhart cautions teachers in using only summative evaluations to make instructional pacing decisions – moving forward whether students mastered the material or not – since this would contradict the purpose of important learning.

One participant adamantly reported that she does not use summative evaluations in her classroom because she uses a combination of assessments. Even though some of her assessments appear to be the traditional type of summative assessments, she explained that she is expected to use the evaluation pieces, but uses non-traditional methods in assessing student work. For example, she views the work as formative learning and establishes an environment where student and teacher conversations are taking place concerning the content of the assessments. Moreover, students are given multiple opportunities to complete and improve their work. Her opinions and practices are supported in literature by Wolf (1993) when he described that summative assessments can be turned into teachable moments. Wolf believes that a thorough review of the data supports the need to check the reliability and validity of the test questions, as well as make instructional decisions about the outcome of learning. These views are parallel to Popham (2008), McMillan (2008), and Brookhart (2009) in how they describe the use of formative assessments which produce similar methods of reflection, analysis, and decision-making.

When discussing summative assessments, the researcher noted that most of the teachers’ responses indicated a lack of understanding that summative assessments may include additional pieces of student work, beyond unit tests and standardized tests. Supported in the literature by Popham (2008) and Taras (2005), summative assessments
also include collections of work over time, such as student portfolios and student project work, which can be assessed using research-based rubrics. This is consistent with the findings of two participants, who use similar approaches in assessing student work through projects, which are evaluated by criterion-based rubrics.

Conclusions

As indicated by the qualitative data analysis, important conclusions were confirmed by the teachers’ discussions of assessments that are supported by literature. It was noted that the participants appeared honest in their answers and were not influenced about whether the use of formative or summative evaluations reflected on their methodology in assessing student work. All agreed that using formative type of assessments, which range from questioning to classwork assignments, provide teachers meaningful information about students’ current level of understanding. Furthermore, the teacher feedback offers valuable information to the students, providing suggestions in which they can improve their overall understanding, resulting in greater success.

All participants shared a similar opinion concerning summative assessments, in which they identify summative assessments as a method to evaluate student mastery; however, all participants believed that student assessments need to consist of a mixture of components. The interpretations are consistent with the literature, supporting that teachers use a variety of assessments components to evaluate and report student work.

Professional Development

Discussion of Findings

As one of the final topics in the qualitative interviews, it was noted by the researcher that the topic of professional development was first met with unenthusiastic
responses. This was largely due to the participants’ overall feeling that most professional development sessions lack meaning and practical ideas. However, focusing on the topic of teacher training with respect to grading practices yielded honest and independent thoughts. Teachers generally agreed that the lack of professional development in the area of grading practices has produced some arbitrary practices.

After comparing and contrasting the results, three notable findings materialized with respect to professional development: (1) teacher training in grading practices; (2) situations in which teachers had conflict with their own grading practices; and (3) the need for professional development in the area of grading practices. The first and third findings are related to one another, supporting the idea that teacher training in grading practices should be the focus of professional development.

In terms of teacher training as it relates to professional development, the researcher noted a surprising result in this study when only one participant received training in grading practices while she was an undergraduate student at college. All of the participants recalled training of some sort in how to develop assessments following curriculum units; however, only one participant was trained in how to grade assessments to validate the grade earned.

The interpretations of this study’s findings with respect to lack of teacher training are consistent with the research conducted by Guskey (2004), Brookhart (1994), and Stiggins and Bridgeford (1985). Collectively, their research findings documented that teachers are ill-equipped in developing and grading assessments that are based on valid measurement standards. Furthermore, Guskey found that limited training and knowledge of grading practices impact the effectiveness of student grades and how it relates to
student achievement, which supports the general conclusion that grades misrepresent student success (Wormeli, 2006). Consistent with the findings of Guskey, Brookhart, and Stiggins and Bridgeford, Ebel and Frisbie (1986) and Tombari and Borich (1999) suggest that technical challenges exist, preventing teachers from accurately measuring student achievement. The recommendations from their studies indicated that teacher preparation courses need to emphasize measurement theory recommendations in order for teachers to develop grading practices based on legitimate principles.

Although the literature supported the participants’ findings with respect to lack of teacher training as undergraduate students, the participants suggested that their current practices in grading have evolved over time due to their teaching experiences, conversations with teacher-mentors, and limited professional development opportunities at their respective schools. This is consistent with Guskey (2004) and Allen’s (2005) findings which concluded that teachers are influenced in their grading practices based on experiences, not measurement courses.

However, Guskey (2004) and Allen’s (2005) findings further indicated that even though some teachers were trained in measurement courses, their practices do not reflect a development of grading that confirms that they were influenced by the measurement recommendations. This is consistent with Brookhart’s 1993 study of 84 classroom teachers, 40 of whom received measurement instruction, who still use a mixture of grading factors to determine student grades, resulting in the belief that measurement instruction was ineffective and teachers tend to develop their own grading practices, which include non-academic factors such as student effort and motivation.
It was an enlightening and honest discussion when the teachers recalled instances in which they changed student grades because they had conflict with their original determination. The participants credit reflection as their greatest ally in changing grades because they want to remain fair and flexible. Even though this topic contributes to the findings of this study, it is not part of the original literature review, and therefore, not supported by literature. However, it is noted that the teachers’ conflicts with their own grading practices were sometimes reported as conflicts with the school’s policies and procedures. This finding is supported in the literature by Austin and McCann (1992) when their study of high school teachers of English and Mathematics reported teachers used varying practices and had difficulty adhering to school guidelines, practices and policies because of the varying degrees of importance teachers placed on academic and non-academic factors.

As the third important finding in this theme, all participants agreed that professional development in the area of grading practices is greatly required in order to understand the impact that teacher decisions have on reporting of student grades. As stated before, the teachers base their current practices on experiences, reflections, and colleagues’ practices; however, they recognized that their practices may be viewed as arbitrary as it relates to measurement recommendations, measures of validity and reliability, and school expectations. Not surprising, the participants’ views are widely supported in the literature.

Literature by Reeves (2011), Allen (2005), Gallagher (1998), and Guskey (2004) collectively supports professional development for teachers to discuss, analyze, and align grading practices that best measure student achievement. Although some schools offer
information about grading practices, as was determined in Austin and McCann’s (1992) study of high school teachers that schools are likely to have informational meetings about grading practices, which included information on: (1) a review of school grading policies; (2) how to compute grades; (3) how to complete report cards; and (4) a presentation of past years’ distribution of grades. The study concluded that the meetings did not include in-depth discussions of grading practices, nor did they have a training component to increase consistency and understanding of measurement standards. Consistent with Austin and McCann’s findings, this would suggest an overall finding that teachers in secondary schools, both middle and high school, identify that even though their respective schools may have mini-sessions on grading practices and measurement standards, they believe that in-depth discussions and training are necessary components to affect change.

**Conclusions**

As indicated by the qualitative data analysis, important conclusions were established as it relates to teacher training and professional development. A lack of teacher training for novice teachers created an environment in which teachers developed their own grading practices with little regard to measurement recommendations. Perhaps one of the reasons teachers have not developed practices that relate to measurement standards is due to the ready-made assessments that accompany textbooks and ancillary materials for each content. Using these pre-made assessments lightens the work for teachers, but they are typically mirrored after the format of standardized tests, resulting in low cognitive levels of student understanding, as supported by Fleming and Chambers
Furthermore, with pre-made assessments, scoring guides are usually included, making the practice of meaningful grading unimportant.

**Limitations**

In research studies, limitations are discussed to analyze possible threats to the study’s validity and to acknowledge existing flaws to the research design. There were several limitations to this study. One limitation was that the school district only allowed the researcher to collect data from the participants during a certain time of the school year. Since the interviews could not take place during most of the second semester, they took place towards the end of the school year. During the last weeks of school, teachers may have felt the stresses of standardized testing and wrapping up the school year, which would have divided their focus. The data may have been different had the collection periods been during a different time of the year, or even throughout the year.

Secondly, a limitation emerged since the researcher is a school administrator within the school division. Although the researcher’s field notes indicated that the participants gave thoughtful and honest answers, the participants may have been reluctant to share all their true thoughts, especially as it relates to their current practices and school expectations. As a third limitation, the participants were all female. This study limited its focus to middle school language arts teachers; therefore, the selected participants only represent one content area within a mid-size school division, from female teachers’ perspectives. The results of this study may be transferable if the range of contexts of this study is realistically replicated in a similar inquiry.
Recommendations for Practice

Based on the results of this study derived from qualitative data and the analysis of the research on grading practices, the following recommendations should be considered as implications for practice. All of the recommendations encourage the participants of this study to become agents of change within their departments by first developing professional learning communities as a department. It is recommended that the learning community form group discussions or book talks that relate to current educational topics such as grading practices. Through the use of journal articles and other literature, the participants can lead fruitful discussions about the positive impact that well-designed grading practices may have on student understanding and overall achievement. The ripple effect may then become a logical step in the reform of establishing grading practices that truly measure student achievement in their one small department as well as other departments within the school.

However, the first logical step within the professional learning community should be focused on conversations. The teachers who participated in this study are recommended to have similar conversations within their language arts departments about grading practices. The participants are encouraged to become leaders within their departments to foster change that represents current research, supporting student-centered assessments as opposed to traditional type assessments which are teacher-centered. To begin this process, it is recommended that the participants lead their departments in meaningful reflection.

Reflection is an important component of teaching. It is recommended that the reflection exercise focus on (1) why teachers grade; (2) what student grades represent in
terms of student achievement; and (3) how they define the primary purposes of grades. This exercise will lead to a greater understanding of the meaning of grades before teachers assign grades. Once each teacher reflects on the purpose of grading, teachers need to share their ideas and opinions with their colleagues within the department in order to create a united understanding of why they grade. This reflection will also assist with communicating the importance of grading to students and their parents. Furthermore, comparing and contrasting their ideas and opinions with their schools’ policies and practices will also generate great discussions that should lead to a greater understanding of the importance of grades.

This study asked teachers to examine their grading practices – their influences, what factors are of most value, and why they chose the factors. It is recommended that the participants lead their language arts departments in a similar examination of their own grading practices to determine if the practices are in alignment with school policies and to determine if student grades truly reflect student achievement. Moreover, it is recommended that the participants from this study lead the departments in reviewing measurement theory recommendations as it relates to grading outlined in research and literature. Contributing to in-depth discussions with other teachers on the topic of academic and non-academic factors is a sensible exercise for practitioners, which may lead to the discovery that their current practice may misrepresent student achievement. It is further suggested that the participants and their colleagues within the department examine student work samples in order to determine if discrepancies exist when determining student grades. It is recommended that an exercise of sharing student work
within the department would lead to a greater understanding of how student performance is measured as practitioners compared to what measurement experts recommend.

This study examined the types of assessments that teachers use to assess student work. The results indicated that the participants need to develop a greater understanding of formative and summative assessments, and how they work together to assess student understanding and mastery. It is recommended that the participants again lead their departments in research to gain a greater understanding of how formative and summative assessments work together to support two main objectives: (1) to assist students in improving their work and level of understanding; and (2) to assist teachers in making sound instructional decisions based on the results of the evaluations. A greater awareness needs to take place by classroom teachers concerning what formative assessments look like in the classroom and what conclusions can be derived from the information they provide. These types of assessments range from questioning techniques to classwork assignments, and they provide valuable and timely information to the teachers to monitor and adjust instruction. In addition, feedback on formative types of assessments also provides instant information to the students in order for them to improve.

With respect to professional development, this study explored teaching training opportunities as they prepared to become teachers as well as current opportunities structured as professional development. It is recommended that the participants of this study communicate to their current administrators the desire to learn more about affecting change in their grading practices that relate to current measurement standards as well as ensuring that grades truly reflect student achievement.
However, classroom teachers are not the only educators who need to participate in some form of professional growth as it relates to developing a greater understanding of grading practices. It is further recommended that the participants initiate conversations with school administrators concerning the topic of current recommendations in grading in order to affect change within the school in developing grading practices that are sensible, realistic, and measure student achievement since the implications for school administrators could include are several. When school administrators examine the grading practices within their own building, the school leaders will gain a better understanding of how student grades are averaged and reported. The results of this examination have implications with teacher evaluation, analysis of student data, and a comprehensive awareness of how student grades are determined.

**Recommendations for Research**

Five recommendations for future research emerged from the findings of this study. First, a similar investigative approach may be applied to other disciplines. This qualitative study was limited to middle school language arts teachers. The findings concluded that all participants shared similar approaches to grading, noting differences in which factors they placed their importance. Future studies which include teachers from content areas of mathematics, science, and social studies may yield varying results. It would be interesting to study the grading practices within the other three disciplines to compare and contrast the findings and conclusions. In the language arts curriculum, the grading is very subjective due to the nature of the student work, yet in the areas of mathematics, science, and social studies, the grading is more content-based and objective,
and the findings and conclusions, as it relates to grading practices, may be similar to one another yet very different from the perspective of the language arts participants.

Secondly, if this investigation were to be replicated, it is recommended that focus groups be used in order to benefit from the discussion among practitioners as they compare and contrast their approaches to grading. In addition, a quantitative investigation which may complement this study could be generated using the results of this study as a survey instrument for similar participants to determine if comparable results exist as it relates to the purposes of grades, which factors teachers use, and how they relate to measurement experts.

As a third recommendation, additional research is needed to determine why a disparity exists between measurement recommendations and practitioners and to determine how to narrow the gap. Although the results of this study concluded that the teachers agree with some of the measurement recommendations, there were a considerable amount of recommendations in which they disagreed, most notably that the teachers of this study do not solely rely on student achievement grades, yet rely on a variety of factors. An investigation of the measurement recommendations may find that the recommendations are outdated and do not relate to the practical use in the classroom.

The fourth recommendation for research is in the area of how to report grades that serve as a multipurpose tool which includes both academic and non-academic factors. Developing a student reporting method that produces a comprehensive picture about student performance, ranging from achievement grades to student behaviors, such as effort, motivation, and responsibility, may alleviate the teachers’ need to average student achievement grades with varying non-achievement factors. This may lead to a
comprehensive understanding by students, parents, and teachers about what student grades represent.

The final recommendation includes future research to determine the effectiveness of teacher training or professional development opportunities to help teachers meet the challenge of grading. With respect to the findings of this study, teachers expressed a great need for professional development in the area of grading practices. Research would determine if the teaching training opportunities were effective or not by examining the teachers’ grading practices, to conclude if teachers followed the guidelines and recommendations by measurement experts, or if they chose to develop grading practices that continued to incorporate a mixed-bag of academic and non-academic factors. The results of the investigation may lead to a restructure of current teacher training programs and professional development opportunities.

In conclusion, the need to continue to examine grading practices within the context of measurement recommendations and assessment best practices may be warranted if teachers continue to report grades that are determined by multiple pieces of information, possibly misrepresenting the accuracy of student grades. Ultimately, for grades to be interpreted with accurate understanding, the grade requires an understanding from both the student receiving the grade and the teacher assigning the grade (Zoeckler, 2007).
List of References
List of References


Fairgrade. (2008). In Fair grades for Fairfax County students. Retrieved from
http://www.fairgrade.org

In W. E. Hathaway (ed.), Testing in the schools: New directions for testing and

Frankel, R. M., & Devers, K. J. (2000). Study design in qualitative research – 1:
Developing questions and assessing resource needs. Education for Health, 13(2),
251-261.

opinions of secondary teachers of academic subjects: Implications for instruction

Merrill/Prentice Hall.

Griswold, P.A. (1993). Beliefs and inferences about grading elicited from student

methodologies? Qualitative approaches to evaluation in education: The silent
scientific revolution, 89-115.

Kappan, 83(10), 775-780.


Guskey, T. (2006). “It Wasn’t Fair!” Educators’ recollections of their experiences as


Educational Testing Service.


Appendices
Appendix A

Interview Guide

Research Questions

1. What influences teachers’ grading practices?
2. What academic factors were considered in determining the grade, and what beliefs and/or values were used to determine those factors?
3. What non-academic factors were considered in determining the grade, and what beliefs and/or values were used to determine those factors?
4. What gaps exist between teachers’ grading practices and recommendations made by measurement experts?

Interview Guide

I. Introduction, purpose of interview and confidentiality

The participants will be informed about the purpose of the interview and the central focus of the questions. They will be assured that the information during the interview will be held in confidence.

Additionally, the participants will be informed that a tape recording will be made of the interview for transcription purposes in order to ensure accuracy of the conversation. During the transcription phase, a pseudonym will be used.

A copy of the transcription will be made available to the interviewees for review, and changes will be made based on their request.

Consent by the interviewees will be granted by signing the consent form.

II. Interview Questions/Topics

1. Demographic Questions
   i. Age
   ii. Gender
iii. Length of teaching experience
iv. Length of teaching experience at current school
v. Length and subject of current teaching assignment

2. Describe your approach to grading.
   i. What factors do you use?
   ii. Why do you use these factors?
   iii. What influences you in using these factors?
   iv. Which do you place the most importance and why?

3. How did you develop your current methods of grading practices?

4. What factors do you classify as academic factors in terms of grading?

5. What factors do you classify as non-academic factors in terms of grading?

6. How do your grading practices compare to current measurement theory? (see Appendix B)

7. Were you trained in the area of developing grading practices that meet measurement theory recommendations? (see Appendix B)

8. Do you ever find that you have conflict with your own grading practices? Give examples.

III. Closure

This now concludes our interview. Is there anything else you would like to add concerning grading practices?

Thank you for your willingness to participate in this qualitative study.
Appendix B

Measurement Theory Recommendations as found in Literature and Research Studies

1. Grading methods should be communicated to students and parents (Brookhart, 2009; Guskey, 2004; McMillan, 2008; O'Connor 2007; Wormeli 2006; Stiggins, Frisbie & Griswold, 1989).

2. Grades should reflect an accurate measure of student’s academic achievement – only factors that contribute to the grade must qualify as achievement (Wormeli, 2006; Winger 2005; Dyrness & Dyrness 2008; Cross & Frary, 1999; Stiggins, Frisbie & Griswold, 1989; Guskey, 2006; O’Connor, 2007).

3. Performance on assessments should match instructional goals and objectives (Brookhart, 2009; McMillan, 2008; Popham 2008).

4. Grading policies and procedures need to be consistent (Brookhart, 2009; McMillan 2008; Marzano, 2000).

5. Student behaviors, effort, and responsibility should not be reflected in student grades (McMillan, 2008; O’Connor, 2007; Winger, 2005, Wormeli, 2006; Stiggins, Frisbie & Griswold, 1989).

6. The purpose of grading is to provide meaningful feedback in terms of comments and not solely scores (Guskey, 2004; McMillan, 2008; Reeves, 2008; Brookhart, 2009).

7. Employ more informal types of assessment on a regular basis – formative type assessments that provide meaningful feedback to help students make informed decisions about how to improve their current learning (Sadler, 1989; Popham 2008; Brookhart 2009; McMillan, 2008).

8. Use summative assessments to provide a final grade as an outcome of learning (Brookhart, 2009; McMillan, 2008).

9. Do not use grades as a motivator – separate academic merit from behavior incentives (Dyrness & Dyrness, 2008).
Appendix C

RESEARCH SUBJECT INFORMATION AND CONSENT FORM

**TITLE:** Grading Practices: Teachers’ Consideration of Academic and Non-Academic Factors

**VCU IRB NO.:** HM13467

This consent form may contain words that you do not understand. Please ask the study staff to explain any words that you do not clearly understand. You may take home an unsigned copy of this consent form to think about or discuss with family or friends before making your decision.

**PURPOSE OF THE STUDY**
The purpose of this study is to examine the academic and non-academic factors teachers consider in their grading practices to gain a better understanding of the discrepancy within teachers’ grading practices. Academic factors include student performance and achievement, while non-academic factors include a variety of items, including, but not limited to: responsibility, effort, attitude, behavior, motivation, attendance, chemistry between teacher and student, class size, and teachers’ efficacy in working with different student populations, such as special education. The results of this study will assist with the understanding of middle school language arts teachers’ approaches to assigning grades in an effort to understand the arbitrary grading practices teachers employ. It will inform practice by bringing awareness to teachers’ understanding of measurement experts’ recommendations, and it will produce meaningful conclusions based on similarities and differences on how teachers determine student grades. Furthermore, the results of this study will help enable additional research for other content areas to determine variation among grading practices and why the variation exists.

**DESCRIPTION OF THE STUDY AND YOUR INVOLVEMENT**
If you decide to be in this research study, you will be asked to sign this consent form after you have had all your questions answered and understand what will happen to you.

In this study you will be asked to participate in one qualitative in-depth audio-taped interview that may last approximately 60 minutes. During the interview, you will be asked questions about your approach to grading; the academic and non-academic factors
you consider while grading; what influences your grading methods; training you
experienced in the area of grading; how you developed your grading practices; if your
grading practices relate to measurement theory, and if you experience conflict within
your own practices. Your interview will be tape recorded to ensure that your responses
are being reported accurately. No names or other identifying details will be recorded on
tape. Significant new findings developed during the course of the research which may
relate to your willingness to continue participation will be provided to you.

RISKS AND DISCOMFORTS
While it is not anticipated that discussing this subject will cause you to be uncomfortable
or fell upset, you do not have to discuss any subjects that you do not wish to and you may
end your participation in the interview at any time.

BENEFITS TO YOU AND OTHERS
You may not get any direct benefit from this study, but, the information we learn from
participants in this study may help us understand and design grading practices that are
standardized with best practices that benefit students and schools.

COSTS
There are no costs for participating in this study other than the time you will spend in the
interview and checking for accuracy once the interview is described.

ALTERNATIVES
You may choose not to participate in this study.

CONFIDENTIALITY
Potentially identifiable information about you will consist of interview notes and
recordings. Data is being collected only for research purposes. Your data will be
identified by ID numbers of pseudonyms, not names, and stored in a locked research
area. All data will be kept in password protected files, and these files will be deleted
upon completion of this project. Paper notes and recordings will be kept in a locked file
cabinet (paper notes) and/or a password protected file (electronic notes) for six months
after the study ends and will be destroyed at that time. Access to all data will be limited
to study personnel.

We will not tell anyone the answers you give us; however, information from the study
and the consent form signed by you may be looked at or copied for research or legal
purposes by Virginia Commonwealth University.

What we find from this study may be presented at meetings or published in papers, but
your name will not ever be used in these presentations or papers.

The interviews will be audio taped, but no names will be recorded. At the beginning of
the session, all members will be asked to use initials only so that no names are recorded.
The tapes and the notes will be stored in a locked cabinet. After the information from the
tapes is typed up, the tapes will be destroyed.
VOLUNTARY PARTICIPATION AND WITHDRAWAL
You do not have to participate in this study. If you choose to participate, you may stop at any time without any penalty. You may also choose not to answer particular questions that are asked in the study.

Your participation in this study may be stopped at any time by the study staff without your consent. The reasons might include:
• you have not followed study instructions;
• administrative reasons require your withdrawal.

QUESTIONS
In the future, you may have questions about your participation in this study. If you have any questions, complaints, or concerns about the research, contact:

Dr. James McMillan, Professor and Director
Foundations of Education
Virginia Commonwealth University
P.O. Box 842020
Richmond, Virginia 23284-2020
Office: 804.827.2620
Fax: 804.828.1323
Email: jmcmillan@vcu.edu

If you have any questions about your rights as a participant in this study, you may contact:

Office for Research
Virginia Commonwealth University
800 East Leigh Street, Suite 113
P.O. Box 980568
Richmond, VA 23298
Telephone: 804-827-2157
CONSENT
I have been given the chance to read this consent form. I understand the information about this study. Questions that I wanted to ask about the study have been answered. My signature says that I am willing to participate in this study. I will receive a copy of the consent form once I have agreed to participate.

<table>
<thead>
<tr>
<th>Participant name printed</th>
<th>Participant signature</th>
<th>Date</th>
</tr>
</thead>
</table>

Name of Person Conducting Informed Consent Discussion/Witness
(Printed)

Signature of Person Conducting Informed Consent Discussion/Witness Date

Principal Investigator Signature (if different from above) Date
Dear Principal:

March 2011

As part of the requirements of Virginia Commonwealth University’s Educational Leadership doctoral program, I am conducting qualitative research for the purpose of analyzing how middle school language arts teachers’ grading practices relate to measurement theory considering both academic and non-academic factors. It is anticipated that teachers representing two middle schools in your mid-size central Virginia school division will participate in this study during the month of June 2011.

This study has been reviewed and approved according to the division policy on research so that I may contact principals within your school district. I am seeking your consent to contact your language arts teachers to request their participation through an in-depth interview, each lasting approximately 30-60 minutes. Participation is strictly voluntary. The promise of strict confidentiality is assured in both the collection and reporting of the data. Any findings obtained in connection with this study will be presented in such a way that no individual school or person will be identifiable. By giving consent and participating in this study, your teachers will be giving me permission to publish aggregated results in my dissertation, in peer reviewed journals, and at professional conferences.

As a school administrator, I am hopeful that the study’s findings will assist with more clearly defining how grading practices relate to measurement theory, and how teachers’ discern the value of academic and non-academic factors. Understanding current assessment beliefs and practices, and formulating relevant professional development aimed at the improvement of teachers’ assessment pedagogies and practices can directly contribute to our students’ educational success.

Should you approve your teachers’ participation in this research study, I will request your assistance with providing the names of your current language arts teachers in your middle school. They will then be contacted by me in another email outlining (a) the purpose of the study, (b) asking for their current practice in regards to percentages of student total grades in comparison of academic and non-academic factors, (c) their willingness to participate, and (d) availability of their time to meet for an interview.
For your convenience, I have attached the interview protocol and consent form for your review. Should you have any questions about this study, please contact me at dyesbeck@hcps.us. Thank you in advance for your time and consideration.

Sincerely,

Diana M. Yesbeck, Assistant Principal
Doctoral Candidate
Virginia Commonwealth University
Appendix E

Letter to Language Arts Teachers

Dear Language Arts Teacher:                 March 2011

As part of the requirements of Virginia Commonwealth University’s Educational Leadership doctoral program, I am conducting qualitative research for the purpose of analyzing how middle school language arts teachers’ grading practices relate to measurement theory considering both academic and non-academic factors. It is anticipated that teachers representing two middle schools in your mid-size central Virginia school division will participate in this study during the month of June 2011.

This study has been reviewed and approved according to the Hanover County Public School policy on research. With this permission and your principal’s consent, I am contacting middle school language arts teachers within your school district. I would welcome your participation through an in-depth audio-taped interview, lasting approximately 60 minutes. Participation is strictly voluntary, and you may withdraw from this study at any time. Your decision whether or not to participate will in no way jeopardize your future relations with your current employer. Please note, that should you determine the need to withdraw from the study at a later date, all data associated with the information you provided will be properly discarded. The promise of strict confidentiality is assured in both the collection and reporting of the data. Any findings obtained in connection with this study will be presented in such a way that no individual school or person will be identifiable. By giving consent and participating in this study, you will be giving me permission to publish aggregated results in my dissertation, in peer reviewed journals, and at professional conferences.

To participate in this study, please answer the two following questions.

1. Generally speaking, what is your current practice of calculating the percent of a student’s grade determined by academic performances as contrasted by the percent determined by non-academic factors? To assist with determining which factors are considered academic and non-academic, the terms and definitions are listed below as found in literature. (For example, 80% of student’s total grade is based on achievement factors only – such as quiz and test scores and 20% of student’s total grade is based on non-academic factors, such as effort, non-graded homework, participation, organization, and attitude.)

2. When are you available during the school week or on weekends to participate in an approximately 30-60-minute audio-taped interview? It is important to conduct
the interview during the last two weeks of the school year based on the administration window of conducting research in your school division.

Once I receive responses from the questions, I will be able to narrow the sample to approximately ten teachers, based on the highest and lowest tenth of teacher responses of the use of strictly academic factors and a combination of both academic and non-academic factors while determining student grades. All teachers who respond to the initial email will be contacted, whether it is to schedule an interview or not.

I am hopeful that the study’s findings will assist with more clearly defining how grading practices relate to measurement theory, and how teachers’ discern the value of academic and non-academic factors. Understanding current assessment beliefs and practices, and formulating relevant professional development aimed at the improvement of teachers’ assessment pedagogies and practices can directly contribute to our students’ educational success. Should you have any questions about this study, please contact me at dyesbeck@hcps.us.

Thank you in advance for your time and willingness to share your approach to grading practices. I would not be able to complete this study without your assistance.

Sincerely,

Diana M. Yesbeck
Doctoral Candidate
Virginia Commonwealth University

**Academic and Non-Academic Factors and Definitions**

**Academic Factors** – those that are considered in grading practices – student achievement or performance which demonstrates mastery of content (Wormeli, 2006).

**Non-Academic Factors** – those that are considered in grading practices – factors that relate to student behaviors, work habits, and attitudes. Examples include effort, participation, organization, non-graded homework, responsibility, motivation, attitude, and attendance (Brookhart, 2009).
Vita

Diana Marie Yesbeck was born in August 1961 in Richmond, Virginia to parents Edward and Vera Yesbeck. She is one of four children, and she and her twin are the middle children of the family. Once she graduated from John Randolph Tucker High School in Richmond, Virginia, she attended Longwood College in Farmville, Virginia, and transferred to Virginia Commonwealth University in Richmond, Virginia after one year. During the next seven years, Diana worked full time at various employers while attending college part time. In March 1987, Diana gave birth to her only child, Jacob Scott Yesbeck. He was her inspiration to complete her undergraduate studies at Virginia Commonwealth University. Much to the delight of Diana’s family, she graduated in May 1988 with a Bachelor of Science Degree in Mathematics from Virginia Commonwealth University.

Diana began teaching in 1990, after attempting a career in Actuarial Science. Her love for teaching children blossomed during the 1990’s. Due to the encouragement of her family, she returned to school to earn a Master of Education Degree from the University of Virginia in January 2004. After serving as a middle school mathematics teacher for twelve successful years, Diana moved into an administrative position at the same middle school. Just two years later, Diana began her doctoral studies at Virginia
Commonwealth University in Educational Leadership and has since become an assistant principal within the same school division.

Diana has presented on numerous occasions throughout her professional years in education on the topics of: (1) Authentic Assessment Practices, 1996 to an audience of educators at Virginia Commonwealth University in Richmond, Virginia, (2) Hanover County’s Approach to Standards of Learning Success, 1999 to an audience of educators and school leaders at the SACS Conference in Williamsburg, Virginia, (3) The Professional Teaching Act, 2002 to an audience of faculty at Chickahominy Middle School, Hanover County, Virginia, and (4) Grading Practices, 2010 – 2011 to audiences of school leaders from Hanover County Public Schools, and faculties at Oak Knoll Middle School and South Anna Elementary School, all within Hanover County, Virginia.