


1-1-2008

The impact of professional development on assessment and grading practices for secondary teachers

Nicole Lynn Roorda
Walden University

Follow this and additional works at: <https://scholarworks.waldenu.edu/dissertations>

 Part of the [Educational Assessment, Evaluation, and Research Commons](#), [Elementary and Middle and Secondary Education Administration Commons](#), and the [Secondary Education and Teaching Commons](#)

This Dissertation is brought to you for free and open access by the Walden Dissertations and Doctoral Studies Collection at ScholarWorks. It has been accepted for inclusion in Walden Dissertations and Doctoral Studies by an authorized administrator of ScholarWorks. For more information, please contact ScholarWorks@waldenu.edu.

Walden University

COLLEGE OF EDUCATION

This is to certify that the doctoral study by

Nicole Roorda

has been found to be complete and satisfactory in all respects,
and that any and all revisions required by
the review committee have been made.

Review Committee

Dr. Howard Carlson, Committee Chairperson, Education Faculty

Dr. Pamela Warrick, Committee Member, Education Faculty

Chief Academic Officer

Denise DeZolt, Ph.D.

Walden University

2008

ABSTRACT

The Impact of Professional Development on Assessment and Grading Practices for
Secondary Teachers

by

Nicole Lynn Roorda

M.A., Regis University, 2003

B.A., Central College, 1994

Doctoral Study Submitted in Partial Fulfillment
of the Requirements for the Degree of
Doctor of Education
Administrator Leadership for Teaching and Learning

Walden University
December 2008

ABSTRACT

Studies suggest that a potential misalignment between assessment and grading practices in reporting secondary student academic achievement has negatively impacted students since grades may not truly reflect actual achievement. Accordingly, the purpose of the current study was to compare secondary teachers' use of academic and nonacademic factors when reporting student achievement before and after professional development on assessment and grading practices. Following change theory, the study investigated the efficacy of employing professional development as a means of influencing more standard and appropriate practices among secondary teachers with regard to using academic and nonacademic factors when employing standards-based reporting to determine grades. The single-group repeated measures design used a random sample of 39 secondary teachers (6th-12th grade) in a nonpublic school in the Midwest who completed online adaptations of the Teacher Survey on Grading Practices (TSGP) and McREL surveys. Chi-square analyses of the TSGP indicated that there was a significant difference in the way teachers used academic and nonacademic factors in determining grades after professional development. Specifically, improvement and mastery, two academic criteria, increased after the intervention, whereas all nonacademic factors on the survey decreased in the amount of weight teachers gave them in determining a summative grade for students. The results of this study led to the conclusion that professional development was an effective means to influence change in grading practices. The study contributes to social change by informing professional development models that promote meaningful conversations about the nature of student achievement, systematic assessment practices, and how to most accurately and equitably assign grades.

The Impact of Professional Development on Assessment and Grading Practices for
Secondary Teachers

by

Nicole Lynn Roorda

M.A., Regis University, 2003

B.A., Central College, 1994

Doctoral Study Submitted in Partial Fulfillment
of the Requirements for the Degree of
Doctor of Education
Administrator Leadership for Teaching and Learning

Walden University

December 2008

UMI Number: H H G I I

Copyright 200J by
Ü [[! ä ä P ä [^ Ä ~ } }

All rights reserved

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI[®]

UMI Microform H H G I I
Copyright 2008 by ProQuest LLC
All rights reserved. This microform edition is protected against
unauthorized copying under Title 17, United States Code.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

DEDICATION

This study is dedicated to David James Roorda, Madeline Rachel Roorda, and Gabrielle Regan Roorda. You are all that I could ever ask for in a family, and I hope that this is an example of what I am always telling you – nothing is out of your reach as long as you dig your heels in and refuse to quit. This is my gift of perseverance for you. I love you.

ACKNOWLEDGMENTS

The journey of completing a doctoral study does not simply involve one person, the researcher. Moreover, the study becomes a labor of love for many people, and without the contributions of all, both large and small, this journey would not have been completed. It is with the utmost gratitude and love that I would like to acknowledge the following people.

First, thank you to my family. I cannot even begin to count the number of times that my family had to sacrifice time as a family unit in order for me to achieve my dream of earning my doctorate. Dave, Maddie, and Gabbie, you will never know the amount of love I have for you that patience that you held while Mommy pursued her dream. I also want to thank my parents, brother and sister (and their families) in helping me achieve this dream.

Secondly, thank you to my doctoral study committee, Dr. Howard Carlson and Dr. Pam Warrick. Thank you for never giving up on me and encouraging me when I became discouraged. Your words of wisdom will always stay with me.

Thank you to Dr. Robert H. Rich, Mid-Continent Research for Education and Learning (McRel), and Ken O'Connor for allowing me to use survey instruments developed by them in order to conduct my study. I appreciate your willingness to share your hard work.

Thank you to my place of employment, which challenges its employees to continue learning about learning. Without the ability to use Test Pilot and pick the brains of some of the smartest people I know, I wouldn't have been able to finish my quest.

Thank you to the students, staff, and administrators of the system I conducted my research within. I am happy that you have been on this journey with me from the inception of the idea. I literally couldn't have done this without you!

Finally, thank you to Cheryl Modlin, my "partner in crime," during our journey at Walden. Cheryl taught me patience and kept me on top of the many demands of being a doctoral student.

TABLE OF CONTENTS

LIST OF TABLES	vi
CHAPTER 1: INTRODUCTION TO THE STUDY	1
Background of the Study	1
Problem Statement.....	4
Nature of the Study.....	5
Purpose of the Study.....	7
Theoretical Framework.....	8
Definitions.....	11
Assumptions.....	12
Scope and Delimitations	13
Limitations	13
Significance of the Study.....	13
Summary and Transition.....	14
CHAPTER 2: LITERATURE REVIEW.....	16
Overview of Assessment and Grading Practices.....	16
The Use of Zeros	24
The Use of Formative and Summative Assessment	26
Emphasis on Developing a Summative Progress Report Based on Most Recent Work.....	27
Separation of Behavior Data (Nonacademic Factors) and Academic Knowledge Data (Academic Factors)	28
Use of Averaging or Points	32
Change Theory	34
CHAPTER 3: RESEARCH METHOD.....	40
Introduction.....	40
Role of the Researcher.....	42
Research and Design Approach	43
Setting and Sample	47
Treatment	49
Instrumentation and Materials.....	50
Data Collection Procedures.....	51
Data Analysis Plan	53
Validity of Study	54
Reliability of Study.....	55
Protection of Participants' Rights	55
Confidentiality.....	56
Risks and Benefits	56

CHAPTER 4: RESULTS.....	58
Questions & Hypotheses.....	58
Data Presentation.....	60
Interpretation.....	68
 CHAPTER 5: SUMMARY, CONCLUSION, AND RECOMMENDATIONS.....	71
Overview.....	71
Interpretation of Findings	74
Implications for Social Change.....	77
Recommendations for Action	78
Recommendations for Further Study.....	80
Conclusion	81
 REFERENCES.....	82
 APPENDIX A: ESTIMATING THE MAGNITUDE OF A CHANGE	88
 APPENDIX B: STANDARDS-BASED GRADING SURVEY	89
 APPENDIX C: INFORMATION SHEET.....	92
 APPENDIX D: CONSENT FORM.....	93
 CURRICULUM VITAE.....	96

LIST OF TABLES

Table 1. Educational Level of Participants in Study	60
Table 2. Frequencies of Responses in Determining Magnitude of Change for Implementing a Standards-Based Reporting System in the Secondary Setting ($n=35$)	62
Table 3. Descriptive Statistics for Comparing Pre/Post Survey Data on Academic and Nonacademic Factors when determining Student Summative Grades ($n=35$)	65

CHAPTER 1: INTRODUCTION TO THE STUDY

Background of the Study

In 1990, the National Education Goals and Indicators were announced by President George H.W. Bush and state governors to address educational issues including school readiness; school completion, student achievement and citizenship; teacher education and development; math and science; adult literacy and lifelong learning; safe, disciplined, and alcohol and drug free schools; and parental participation (National Education Goals Panel, 1991). The goals and indicators were to be met by the year 2000.

As a result of the federal legislation, states were required to develop standards and benchmarks for student sin their public school systems. One of the expected outcomes from the law were to align federal programming with curriculum, instruction, and assessment. As students attained these standards and benchmarks, they would also meet the National Education Goal Indicators for student achievement and citizenship. Iowa was the only state in the United States that chose to have each of its 559 school districts develop their own standards and benchmarks, rather than opting to adopt state standards and benchmarks.

One area that measures student performance, as mandated in the No Child Left Behind (NCLB) Act of 2001 (Public Law 107-110: 2002), relates to grading and assessment practices. This mandate has not received the same level of scrutiny that other mandates have. A study of a nonpublic school system that focused on student achievement grades revealed a discrepancy regarding how grades are derived across the

system. To combat this inequity, a reporting system was established to examine the systemic issue eight year ago (B. Lindahl, personal communication, July 2005).

Currently, the reporting system committee collects information on assessment and grading practices (dependent variable) used by 6th - 12th grade teachers in the system before making a recommendation to the administrators for a 6th – 12th grade reporting pilot. The pilot will be directed by collecting data before and after professional development (independent variable) on assessment and grading practices based on the work of O'Connor (2002, 2004).

The professional development sessions will include, but are not limited to (a) discussion of the magnitude of change involved with the implementation of a standards-based reporting system, (b) analysis of current practices around assessment and grading practices, (c) training on assessment and grading practices based on the work of a variety of researchers (Brookhart, 1993; Buisck, 2000; Guskey& Bailey, 2001; Marzano, 2003; Marzano, 2006; O'Connor, 2002; O'Connor, 2004; Stiggins, 1997; Tomlinson & McTighe, 2006; Wiggins, 1994; Wormeli, 2006), and (d) horizontal and vertical grade level discussions on the consistent use of assessment and grading practices.

O'Connor (2002) identified eight assessment and grading practice guidelines that often spark essential dialogue when a school district is considering a standards-based reporting system (p. 46). Based on studies by this author (Lindahl & Roorda, 2006), the secondary school teaching staff of the nonpublic school system in the Midwest is in conflict with O'Connor's guidelines regarding current assessment and grading practices.

These conflicts may result in students receiving grades that may not be a true reflection of their academic achievements.

A study examining secondary school teachers' use of academic and nonacademic factors (dependent variable) when determining students' grades will benefit the body of evidence being gathered around assessment and grading practices in the secondary setting for several reasons. By analyzing the use of academic and nonacademic factors when determining student grades before and after professional development (independent variable) on assessment and grading practices, a heightened awareness of the actual components of a final grade will be made available to stakeholders within the system. By comparing the perceived components of a grade, the case for implementing a standards-based reporting system and its associated assessment and grading practices may be built. The data collected from this study could be used to expand research on the actual achievement data colleges are analyzing to determine students' admissions to postsecondary educational institutions. Finally, the data collected from this study will provide insight into the purposes for assessment and grading in secondary schools in the United States.

There are many possible factors contributing to this problem, including (a) assigning extra credit to students, (b) assessing and grading students with the inclusion of nonacademic factors, and (c) the differences in assessment and grading practices for students that occur within a department or school setting. This study will contribute to the body of knowledge needed to address this problem by looking at the assessment and grading practices used by a sample of secondary school teachers to determine student

grades. The study will also contribute to the body of knowledge analyzing differences in assessment and grading practices for different content areas.

Problem Statement

There is a problem in assessment and grading practices as they relate to determining students' grades in a secondary setting. That problem, specifically, is a misalignment between assessment and grading practices and the reporting of academic achievement (Carr & Artman, 2002; Flynn, Mesibov, Vermette, & Smith, 2004; Reeves, 2002; Squires, 2005). Currently, both academic (student achievement) factors and nonacademic (homework, participation, behaviors) contribute to the determination of summative grades in many classrooms (Lindahl & Roorda, 2006). However, if teachers are reporting academic achievement progress, nonacademic factors should not be reflected in students' grades. When nonacademic material is a factor, student grades may not be a true reflection of actual academic achievement.

While the use of assessment and grading practices in determining students' grades continues to be a topic of conversation nationally, the availability of data on the use of academic and nonacademic factors when determining students' grades is critical. Past studies have shown that assessments and grades are used for a variety of reasons in public schools: (a) to sort students, (b) to classify students, (c) provide scholarships, and (d) even track students (Chappus, Stiggins, Arter, & Chappus, 2004; Elliott, 2005; Guskey & Bailey, 2001; Reeves, 2004a; Reeves, 2004b). What is not known is the impact assessment and grading practices have on student grades in nonpublic entities that have a

standards-based reporting system, and whether the use of academic and nonacademic factors in assessment and grading practices are impacted by professional development.

Nature of the Study

During this study, the researcher utilized an online testing survey instrument called Test Pilot to collect data regarding how teachers perceive the magnitude of change associated with the implementation of a standards-based reporting system in the secondary setting (dependent variable). The survey was also used to collect data on the use of assessment and grading practices both before and after professional development sessions (independent variable) were provide on assessment and grading practices. This study addressed the following questions:

Question 1: With the implementation of a standards-based reporting system and its associated assessment and grading practices, is there a significant difference in the perception of nonpublic secondary teachers about the magnitude of change involved?

The null hypothesis for this question is (H_0): With the implementation of a standards-based reporting system and its associated assessment and grading practices, there is no significant difference in the perception of nonpublic secondary teachers about the magnitude of change involved.

The alternative hypothesis for this question is (H_1): With the implementation of a standards-based reporting system and its associated assessment and grading practices, there is a significant difference in the perception of nonpublic secondary teachers about the magnitude of change involved.

Independent variable - implementation a standards-based reporting system and its associated assessment and grading practices

Dependent variable - perception of nonpublic secondary teachers

Question 2: After receiving professional development about assessment and grading practices in a standards-based reporting system, is there a significant difference in the assessment and grading practices of secondary teachers with regards to academic and nonacademic factors when determining students' grades?

The null hypothesis for this question is (H_0): After receiving professional development about assessment and grading practices in a standards-based reporting system, there is no significant difference in the assessment and grading practices of secondary teachers with regards to academic and nonacademic factors when determining students' grades.

The alternative hypothesis for this question is (H_1): After receiving professional development about assessment and grading practices in a standards-based reporting system, there is a significant difference in the assessment and grading practices of secondary teachers with regards to academic and nonacademic factors when determining students' grades.

Independent variable – professional development

Dependent variable – assessment and grading practices

Question 3: Is there a significant correlation between the teachers' content area taught and their assessment and grading practices?

The null hypothesis for this question is (H_0): There is no significant correlation between the teachers' content area taught and their assessment and grading practices.

The alternative hypothesis for this question is (H_1): There is a significant correlation between the teachers' content area taught and their assessment and grading practices.

Independent variable – teachers' content area taught

Dependent variable – assessment and grading practices

Purpose of the Study

The purpose of this repeated measures quantitative study is to compare secondary teachers' use of academic and nonacademic factors when reporting student achievement before and after professional development on assessment and grading practices in a standards-based reporting system. This study included secondary school teachers (6th – 12th grade) in a nonpublic school system in the Midwest who completed the Teacher Survey on Grading Practices, a previously validated pre- and posttest survey developed by and used with permission of Rich (2001). This survey identified the use of academic and nonacademic factors when determining summative student report card grades. In addition, a validated frequency count survey designed by and used with permission from Mid-Continent Research for Education and Learning (McREL: 2006) provided data on the magnitude of change for an impending initiative, in this case, a standards-based reporting system with its associated assessment and grading practices.

When using the Teacher Survey on Grading Practices, a Likert scale survey developed and validated by Rich (2001), teachers will complete the survey using an

online survey vehicle. For the purpose of this study, academic factors were defined as those indicators that are directly tied to student academic achievement including subject-specific content learned, reasoning and thinking skills demonstrated, and communication skills demonstrated. Nonacademic achievement factors are defined as indicators that could be included in a grade such as behavior, attendance, participation, and work completion. Teachers also answered several questions that led to the magnitude of change associated with the implementation of a standards-based reporting system, and its associated assessment and grading practices.

Theoretical Framework

Lewin, (1947), is widely regarded as the founder of modern social psychology (Clark, 200) and coined the term *change theory*. Within this theory, Lewin proposed that organizations were unfrozen, changed, and then refroze. The key to the change process, whether it was at an individual or group level involved “painful unlearning without loss of ego identify and difficult relearning as one cognitively attempted to restructure one’s thoughts, perceptions, feelings, and attitudes” (Schein, n.d., p.1). Further, Lewin felt there were two ways to effect change within an organization. One involved forcing change on individuals, while holding them accountable for change, while the other involved “removing the restraining forces that inhibit people from understanding or being successful with the proposed change” (McREL, 2006, p.26).

Cuban (1996) took change theory a step further by defining change as either being incremental or fundamental. Incremental change is to “improve efficiency and effectiveness in existing structures or schooling, including teacher” (p.76). Cuban used a

car analogy to make the point of an incremental change by stating that “incremental change is like adding new tires, brakes, battery, and a water pump to a car (p. 76).

Fundamental change, as described by Cuban (p. 76) “transforms and alters, permanently, the basic structural framework of the system.” Extending the car analogy, fundamental change would be overhauling the “old jalopy that is beyond repair” (p.77).

Waters, McNulty, and Marzano (2004a, 2004b, 2005) examined the type of leadership associated with both incremental and fundamental changes. They determined that “some innovations require changes that are gradual and subtle,” (p. 49) or first order change, while others “require changes that are drastic and dramatic” (p. 49) or second order change. The terms first and second order change were used in the survey for this study.

First order change is “the next most obvious step to take in a school or district” (Marzano et al. 2005, p.66). These changes do not vary too drastically from the existing norms within an organization and align with the individual’s personal beliefs and values. Second order changes, conversely, seem to make an apparent break from past practices, require a new skill or knowledge set to be acquired, and may not align closely with an individual’s personal belief and value system. Marzano et al. defined this change as deep change that “alters the system in a fundamental way, offering a dramatic shift in direction and requiring new ways of thinking and acting” (p. 66). During the course of the study, the data collected impacted the action planning that administrators would follow in implementing a standards-based reporting system in the secondary system.

The research associated with first and second order change will be applied to this study through the use of a frequency count to determine the magnitude of change perceived by the secondary staff in regards to the implementation of a standards-based reporting system in the secondary setting.

When teachers assign a summative grade to a student, the weight given to academic factors and nonacademic factors may not provide a valid reflection of the student's achievement. The theory of validity was developed by Messick (1989) and has informed the study of validity in measurement. This theory holds that when, "test scores are interpreted and used" (p. 5) they should be analyzed for two aspects of validity: the intended function of the score (interpretation or use) and the source of justification (empirical evidence or social consequences). Messick took the two sources of validity, interpretation and empirical evidence, and crossed them in a matrix with four representative quadrants: (a) construct validity, (b) relevance and utility, (c) value and implications, and (d) social consequences. Messick contended that validity is "a unified concept based on the concept of construct validity" (p. 8). Brookhart (1993) has made connections between Messick's theory of validity and grading practices used by teachers. Applying this theory to the present study, it is hypothesized that there will be statistically significant difference in the use of academic and nonacademic factors contributing to students' grades after participation in professional development on assessment and grading practices in a standards-based system.

In analyzing the data that are retrieved during the course of this study, the researcher will attempt to determine the magnitude of change associated with

implementing a standards-based reporting system, based on the definition so first and second order change developed by Waters et al. (2004a, 2004b). Second, data were analyzed to determine support or negation of Lewin's (1947) theory of change

Definition of Terms

There are several key terms in this study including:

Academic Factors: indicators that are directly tied to student academic achievement, including subject-specific content learned, reasoning and thinking skills demonstrated, and communication skills demonstrated.

Assessment Practices: a way to provide feedback to students on their performance on classroom related tasks. The purpose of providing assessment is to guide both student and teacher towards better learning and teaching practices.

Formative Assessments: ongoing, regular checking of students' understanding (checkpoints). Formative assessments provide risk-free initial attempts to practice the concepts presented. These assessments are designed to provide direction for improvement and/or adjustment to an educational program for individual students or the whole class.

Grading Practices: methods used to determine student achievement in a class, including but not limited to, using points or percentages on assignments, assigning zeros for late work, and including nonacademic factors such as class participation or late work when determining a student's final grade.

Nonacademic Factors: indicators that could be included in a grade such as behavior, attendance, participation, and work completion.

Standards-Based Grading Practices: the practices used by teachers to report student's progress towards meeting district determined standards and benchmarks.

Summative Assessments: used to assess students' achievement at the end of the instructional period of time (mile posts). A summative grade might include only formative information or a compilation of summative marks during a semester, for example.

Assumptions

The core assumptions made by the researcher are:

1. The participants are willing participants in the study.
2. The role of the researcher as an employee of the Acme Education Corporation will not inhibit teachers' participation in the study.
3. The sample setting is representative of similar secondary nonpublic schools in the Midwest.
4. The participants will provide honest data on both the presurvey and postsurvey.

A threat to validity may be the exposure all teachers have had in regard to the effects of different grading practices on students' grades. In addition, there is no way to measure the truthfulness of the responses provided by respondents at any level.

Scope and Delimitations

The independent variable in this study will be professional development sessions on assessment and grading practices that the secondary teachers will participate in during the course of the study. The researcher will compare the data gathered about perceptions of the staff around student assessment and grading practices, the dependent variable, before and after participating in professional development revolving around assessment and grading practices in a standards-based reporting system. An online survey will be administered to collect and analyze data. This study was delimited to the secondary teachers of one selected nonpublic school system in the Midwest, representing grades 6 – 12.

Limitations

The findings of this study will be based on results of a pre-validated survey constructed by researcher Rich (2001) as well as a survey designed by and used with permission of McREL (2006) on the magnitude of change for the implementation of a standards-based reporting system and its associated assessment and grading practices. Therefore, the results may relate to this system along and may not be generalized to other educational settings.

Significance of the Study

The body of evidence around assessment and grading practices will be augmented based on data collected in the study. Stakeholders within a nonpublic school system will receive data that may heighten the awareness of the components of students' grades. The data collected may build the case for a standards-based reporting system in the secondary

setting. Colleges could be solicited to determine how they use the achievement data provided to them for admissions to postsecondary academic institutions. Finally, assessment and grading practices within the secondary setting of American academic institutions could be examined.

Summary and Transition

The use of assessment and grading practices in determining students' grades occurs daily in secondary settings nationwide. The data that are used to determine student grades come from various sources, both academic and nonacademic. Past research has been completed on the purposes of grading within the secondary setting, but there is no research on the impact assessment and grading practices have on student grades in nonpublic entities that have a standards-based reporting system. This study also investigated whether the use of academic and nonacademic factors in assessment and grading practices were impacted by professional development.

Chapter 2 of this study describes the literature base for the three research questions found within this study and how the literature was found. An examination of the history of grading and assessment practices is examined, as well as the research behind common assessment and grading practices, like the use of zeros and utilizing nonacademic factors like class participation and attendance when determining grades. The literature associated with change theory is also explored in this chapter.

Chapter 3 justifies the use of the repeated measures quantitative design of the study. The chapter considers why quantitative research was an effective method for conducting research on teacher perceptions before and after professional development.

The role of the researcher, research design and methodology, treatment and instrumentation and materials were also described. Data collection procedures and analysis are also described in this chapter.

Chapter 4 is focused on the research questions and hypotheses and their results. A complete data analysis is given, with tables that relate to the findings. Chapter 4 concludes with a summarization and interpretation of the data examined.

Chapter 5 summarizes the purpose of the study and the methodology used to investigate the research question. In addition, the chapter discusses the interpretation of data collected as it relates to current literature. Implications for social change and recommendations for further study are also provided.

CHAPTER 2:

LITERATURE REVIEW

Overview of Assessment and Grading Practices

To examine both academic and nonacademic factors as they relate to assessment and grading practices in a standards-based reporting system, the researcher reviewed research on several broad topics, including the theory of validity, developed by Messick (1989), and change theory, developed by Lewin (1947). Literature on the historical context of assessment and grading practices and how those were used and reported to stakeholders in conjunction with the research behind assessment and grading practices as described by O'Connor (2002, 2004) provide an important overview of the process. Changing assessment and grading practices can have a major impact on the culture of a school district and its stakeholders.

In determining which literature would be reviewed for this study, the researcher relied on the use of the EBSCO research database, review of primary source journals and text, and review of dissertations and doctoral studies that had similar areas of focus. The theoretical backdrop of this study is that when teachers assigned a summative grade to a student, the weight given to academic factors and nonacademic factors may not provide a valid reflection of the student's achievement. The theory of validity was developed by Messick (1989) informed the study of validity in measurement. This theory contends that when "test scores are interpreted and used" (p. 5) they should be analyzed for two aspects of validity. In his theory, Messick contended that validity is "a unified concept based on the concept of construct validity" (p. 8). Brookhart (1993) made connections between

Messick's theory of validity and grading practices used by teachers. Brookhart contended that by applying Messick's framework to classroom assessment, teachers had a very clear way to measure student progress. All assessment should lead to further student learning and should also be void of unintended consequences.

Many researchers have analyzed teachers' perceptions about what is included when marking a student's final grade (Agnew, 1985; Johnson, 2001; Wiggins, 1994). Teachers indicated that behavioral issues were combined with academic achievement when reporting the summative grade. These factors, coupled with the teacher or district's position on the use of zeroes and allowing students to retake assignments and assessment have a profound impact on student grades. It has been noted that these practices and beliefs may not be consistent within a building, let alone a district. Marzano (2000) noted that "a single letter grade or percentage score is not a good way to report student achievement in any subject area because it simply cannot present the level of detailed feedback necessary for effective learning" (p. 106). Marzano suggested that grading practices are more reflective of student achievement towards district or state developed standards and benchmarks for courses.

In 1983 the National Commission on Excellence in Education published the report *A Nation at Risk*. Researchers argued that the United States had become complacent regarding education. The report's authors urged citizens to take a hard look at the rigor of the curriculum that was presented to students, due to the fact that other nations' students were beginning to close the achievement gap. From this report, a new movement arose in educational practices in the form of standards-based learning.

States (or local districts) developed content standards to encourage the highest achievement of each and every student by defining the knowledge, concepts, and skills that students should acquire at each grade level. States were also asked to provide evidence of an aligned assessment that measured the content standards. With the advent of NCLB achievement data was reported to the public. This information was used to determine whether a school (or district) was in need of improvement in accordance with NCLB legislation (Public Law 107-110). While many initiatives have come forth to support and enhance the implementation of the initial standards that were created, three areas have seemingly remained stagnant in most schools: (a) assessment, (b) grading practices, and (c) report cards. Although neither NCLB legislation, nor its predecessor The Coleman Report (1966) specifically called for a reform in grading practices across the country, each individually called for educational reform that impacted the way student achievement was recorded and reported back to the general public.

As school districts determined the best way to communicate academic achievement to stakeholders, grades became a point of controversy. What have become "...one of the most sacred traditions in American education" (Olson, 1995, p. 24) have been scrutinized as to whether or not they authentically provide stakeholders the clearest view of a student's academic achievement. "If grading and reporting do not relate grades, back to standards, they are giving a mixed message. Our grading practices must reflect and illuminate those standards" (Busick, 2000, p. 73). Two questions that have been asked across the nation are, what (if any) purpose do grades serve in a standards-based reporting system? When there were no standards-based reporting, grades were a

vital instrument to determine student success. Test scores are now easily compiled and compared, and the question arises regarding how assessment and grading practices play into the reporting of student achievement on a summative student report card.

Teachers have used grades since the 1800s and researchers such as Airasian (1994) have explained that teachers use assessment and grades for five primary reasons.

1. Administrative purposes.
2. To give students feedback about their progress and achievement.
3. To provide guidance to students about future course work
4. To provide guidance to teachers for instructional planning.
5. To motivate students.

“Consciously or not, a teacher’s beliefs and perceptions about life in general influence his or her teaching approach, expectations about student learning, and how he or she goes about grading students” (Carr & Farr, 2000, p. 45). Assessment and grading practices are very personal and can differ greatly from district to district, school to school, and ultimately, classroom to classroom.

Kohn (1993) suggested that there are three main purposes for grading: giving feedback, motivating, and sorting. He added “...grades in particular undermined intrinsic motivation and learning, which only serves to increase our reliance on them” (p. 201). Writing on the unintended negative consequences of grading, Kohn contended that “teachers and parents who care about learning need to do everything in their power to help students forget that grades exist” (p. 206). Kohn cited studies that show that teachers may give a particular assignment two different grades if the work is submitted

on two separate occasions. This variation is greater when multiple teachers evaluate a piece of work on two separate occasions.

Some researchers contend that teacher collaboration increases student achievement and could alleviate some of these grading differences. In their book, *Whatever It Takes: How Professional Learning Communities Respond When Kids Don't Learn*, DuFour, DuFour, Eaker, and Karhanek (2004) suggested "Schools can monitor the learning of each student on a timely basis...if teachers work together to develop common assessments, analyze the results, and assist one another with areas of concern" (p.174). The importance of the development of common practices for assessment is echoed by Chappuis, Stiggins, Arter, and Chappuis (2004) which reminded teachers that

A balanced assessment system takes advantage of assessment OF learning and assessment FOR learning; each can make essential contributions. When both are present in the system, assessment becomes more than just an index of school success. It also serves at the cause of that success. (p. 25)

Even when faced with vast discrepancy regarding the issuing of grades, most parents are not willing to consider a world without grades. According to Kohn (1993), "...one reason so many [parents] seem obsessed with their children's grades and test scores is that this may be their only window into what happens at school" (p. 210). One of the arguments against removing grades from the middle school, and in particular, high school setting, was that the absence of grades could hinder student admission to college.

In personal communications with the deans of admissions at both Harvard and Brown Universities, Kohn discovered that students who have nontraditional report cards often receive more opportunities for enrollment and financial assistance than students

with traditional report cards because admissions officers have to spend a greater amount of time looking at the actual credentials.

While this notion may appease parents, another large contingency of stakeholders still has to be considered and that is the teachers. Studies have shown that teachers do not want to change their grading practices. One frequently cited reason is the perception that grades served the purpose of holding students accountable for their studies. For most teachers, the thought of altering grading practices or eliminating grades altogether closed down the conversation in regards to standards-based reporting.

Kohn (1993) advocated for the elimination of grades, however, he did offer alternative strategies for teachers to use in order to minimize the stigmatism associated with grades and grading practices. The strategies included:

1. Limit the number of assignments for which you give a letter or number grade, or better yet, stop the practice altogether.
2. Limit the number of gradients. For example, switch from A/B/C/D/F to check-plus/check/check-minus.
3. Reduce the number of possible grades to two: A and incomplete.
4. Never grade students while they are still learning something.
5. Never grade for effort.
6. Never grade on a curve.
7. Bring students in on the evaluation process. (pp. 208-209)

While other educational researchers have not fully embraced all seven of these strategies, there is support for some of the strategies presented. Marzano (2000) argued that academic achievement should be the primary factor when developing a grade, but he also noted that a compromise could be reached to include behavior as part of a grade. He contended, that while "...it is appropriate to provide feedback to students on their effort, behavior, and attendance, ideally this feedback should be kept separate from that provided on academic achievement" (p. 39).

O'Connor (2002) and Stiggins (1997) both agreed with Kohn in the fact that effort should not be graded. They went as far as to add participation and late work into this non-gradable mix. O'Connor stated, "Strong effort, active participation, and positive attitude are highly valued attributes, but they are reporting variables, not grading variables" (p. 100). Stiggins added that in terms of effort "...definitions of trying hard vary greatly from teacher to teacher" (p. 418). He also noted that in terms of participation, a student's personality may come into play more than their knowledge base. O'Connor noted

To a considerable extent, personal and social characteristics do contribute to achievement, but including a mark for attitude as part of a mark for product blurs the assessment of the product and affects the validity and thus the meaning of the grade. Also, including a mark for effort of any of these characteristics means a double benefit for successful students and a double (or triple or quadruple) jeopardy for less successful students. This is clearly unfair. (p. 72)

Trumble (2000) argued that not every piece of work assigned in a classroom must be graded. "If student have not mastered something but still honing a skill, a teacher can wait to grade them" (p. 37). This quotation reinforces the thinking of other researchers

that investigate standards-based reporting assessment and grading practices like the one mentioned previously. However, data in the study indicated that this would be a shift in thinking for secondary teachers in the nonpublic system.

For some teachers this could constitute a philosophical change in the way grading and assessment practices are used within their classroom, school, or district. Change is difficult, especially when the change means abandoning practices that are almost engrained in our day-to-day functions. By educating our stakeholders on the historical perspective on why a change may be necessary, valuing their questions, anxieties, and insights, and making this decision systemic, there can be substantial changes within a district in relation to their grading reporting practices. The key to remember when working through this, or any change process, is that change takes time, and that each and every stakeholder brings knowledge and insights with them that will help make informed decisions for students.

Researcher O'Connor identified eight grading guidelines that often spark essential dialogue when examining assessment and grading practices (O'Connor, 2002, p. 46).

O'Connor's guidelines include:

1. The use of zeros.
2. The use of formative and summative assessment.
3. Emphasis on developing a summative progress report grade based on most recent work.
4. Separation of behavior data and academic knowledge data.
5. The use of averaging or points when determining a student's grade.

6. Multiple opportunities for students to complete assessments.
7. Students' involvement in the assessment and grading process.
8. The use of a quality record keeping system (management) to determine students' summative grades. (pp. 62-160)

The Use of Zeros

Research analyzed the use of a zero for grading purposes was a common practice in both middle and high school. Teachers often used zeros to indicate students' incompleteness of assignment(s). The zeros often are detrimental to students' grades. Once students receive multiple zeros, it is often difficult to recover. Zeros have traditionally been used in classrooms as a form of accountability for students (Canady & Hotchkis, 1989; Guskey, 2000; Guskey & Bailey, 2001; Marzano, 2000; O'Connor, 2002). This grading practice has been used in schools based on the assumption that by assigning a zero on an assignment that was not turned in, the student will take more responsibility for completing work assigned in the future. According to Guskey (2000), "No studies support low grades or marks as punishments. Instead of prompting greater effort, low grades more often cause students to withdraw from learning" (p. 25).

Several researchers (Carr & Farr, 2000; O'Connor, 2002; Trumbull, 2000) offered that redefining the mark given to students who fail to turn in assignments might be considered in a standards-based reporting system. Rather than a zero, teachers might assign the student a grade of incomplete, knowing the student had to either complete the assignment that was not turned in or complete another activity that would show attainment of the concepts assessed by the assignment. Before developing a district

policy on the use of zeros and the responsibilities of students and staff for work completion, districts should be advised to look at resource and space allocation.

One question that could be reflected upon when reviewing assessment and grading practices is whether the zeroes are a reflection of the student's academic knowledge, or more a reflection of the student's work habits that were pleasing to the teacher? Did the completion of academic work reflect the student's knowledge of the academic area? This is a scenario that is often played out to the extremes of the learning spectrum with at risk students and talented and gifted students. At risk students often do not turn work in, not because they cannot complete the work, rather, they simply do not see the value of the assignment given to them. Students who possess vast knowledge in a content area sometimes fail a class because their grade is based on behaviors outside their academic knowledge.

Zeros have traditionally been used in classrooms as a form of accountability for students. If a student failed to turn in an assignment, the teacher assigned the student no points for the assignment. This procedure has been used in schools based on the assumption that by assigning a student a zero on an assignment that was not turned in, the student would take more responsibility for completing work assigned in the future. According to Guskey (2000), "No studies support low grades or marks as punishments. Instead of prompting greater effort, low grades more often cause students to withdraw from learning" (p.25).

In previous research completed by the researcher (Lindahl & Roorda, 2006), it was noted that teachers reported using zeroes in their classrooms as a form of

punishment. Further research could be completed to investigate how the use of zeroes impacted individual student work when determining their final grade.

The discrepancy in points between a failing grade and a zero should also be discussed (O'Connor, 2002, p. 151). The question often discussed revolves around if a student fails to turn in his/her work, is it ethical to assign him/her a grade that, based on a 100 point assignment, can put him/her 59 points below the cutoff for a failing grade?

An alternative offered by several researchers (Carr & Farr, 2000; O'Connor, 2002; Turnbull, 2000) is redefining the mark given to students who fail to turn in assignments. Rather than a zero, one alternative would be to assign the student a grade of incomplete, with the student being required to either complete the assignment that was not turned in for the class or complete another activity that would show attainment of the concepts covered by the assignment. The concept of implementing the use of an incomplete within a district should not be undertaken without reviewing the financial and time allocation available in the district to provide extra opportunities for students to complete incomplete work.

The Use of Formative and Summative Assessment

In *The Learning Leader: How to Focus School Improvement for Better Results* (2006), Reeves discussed the importance of using formative assessment to guide instruction. He stated

To be effective, the frequent common assessments used by the most successful schools are not isolated events but integral parts of the teaching, leadership, and learning cycle. Assessment informs teaching; leadership provides the time and resources for teachers to respond to assessment results; and students use assessment feedback as a series of cues for improved performance. (p.87)

Zemelman, Daniels, and Hyde (1998) concurred with Reeves (2006) about the importance of using formative assessments in the classroom to drive instruction; however, they further contended the summative evaluation should be used sparingly, if at all. “Summative evaluation doesn’t aim to nurture learning at all, but merely quantifies what has been learned up to a given point. It isn’t educational; it’s just a way of reporting periodically to outsiders about what has been studied or learned” (Zemelman et al., p.247).

The editors of the book *A Taxonomy for Learning, Teaching and Assessing: A Revision of Bloom’s Taxonomy of Educational Objectives* commented on the fact that “summative assessment provides the data teachers need to make and justify the grades they assign students,” while formative assessments are used primarily to “guide instructional decisions” (Anderson et al., 2001, p. 246).

In previous research conducted with secondary teachers (Lindahl & Roorda, 2006), only 43% of the teachers surveyed noted that they used both formative and summative assessment in their assessment and grading practices.

Emphasis on Developing a Summative Progress Report Based on Most Recent Work

Several researchers noted that students’ summative grade reports should reflect the most consistent and current work they have done in progressing toward standards and

benchmarks (Guskey, 1996; Guskey & Bailey, 2001; Hart, 1996; O'Connor, 2004).

Students who started slowly in class but gained steady progression often were unable to attain a *good* grade due to the fact they had struggled with the concepts presented at the beginning of the course. The summative grade often times was not a true reflection of the student's actual knowledge attainment of the topic area. Guskey and Bailey (2001) remind us, "Teachers should base grades on the most consistent level of performance, not the whole range of performance" (p. 114).

Previous research conducted by the author (Lindahl & Roorda, 2006) noted that in a secondary setting, only 2 of 23 teachers (9%) indicated they used most recent work to determine a student's summative report grade. Rather, the majority (91%) indicated the use of averaging grades to determine a student's summative grade. This data is consistent with the work of other researchers, and would indicate a potential misalignment in the implementation of a standards-based reporting system.

Separation of Behavior Data (Nonacademic Factors) and Academic Knowledge Data (Academic Factors)

Grading practices are subjective and can vary from classroom to classroom within a school, thus provide little consistency for students. Some of the factors that teachers weigh into grades are effort, attitude, and achievement (Carr & Harris, 2001; O'Connor, 2002; O'Shea, 2005; Stiggins, 1997; Trumbull, 2000; Wiggins, 1994). Students in a classroom could receive the same grade for various reasons. Several researchers (O'Connor, 2002; Stiggins, 1997; Tomlinson & McTighe, 2006) suggested providing separate grades for academics and social behaviors. When the academic and social

behaviors of students are separated, a more accurate vision of the whole child is given.

The separation of grades into various components can also help staffs look for trends in students with regard to their social behaviors, academic behaviors, and work habits.

Gronlund and Linn, as quoted in O'Connor, 2002 stated

Letter grades are likely to be most meaningful and useful when they represent achievement only. If they are contaminated by such extraneous factors as effort, the amount of work completed (rather than the quality of the work), personal conduct and so on, their interpretation will be hopelessly confused. When letter grades combine various aspects of pupil development, not only do they lose their meaning as a measure of achievement, but they also suppress information concerning other important aspects of development. (p.88)

Stiggins (1997) offered that some of the areas that are included in grades are very hard to measure. While one student has a very outgoing personality, does this mean that this student has a better grasp of the academic content than the student who quietly listens in class, but does not participate as much? In some classes attitude is also a factor when determining grades. Stiggins (1997) argued the measurability of this trait, and also questioned the subjectivity of factoring this characteristic into a student's grade. One teacher may frown upon what is a pleasing personality to another teacher within a school.

A majority of the studies looking at this topic involve the use of a teacher completed survey (Agnew, 1985; Johnson, 2001; Wiggins, 1994). In these studies, teachers indicated that effort, behavior, work habits, external attributes like socioeconomic status and language, as well as teacher expectations for students were combined with academic achievement when reporting the summative grade. These factors, coupled with the teacher's or district's position on the use of zeroes and allowing students to retake assignments and assessments can have a profound impact on the grade

of a student. It has been noted that those practices and beliefs may not be consistent within a building, let alone a district.

Researcher Rich (2001) took the research a step further. Similar to his colleagues, Rich surveyed teachers using a Likert scale (1-4, with 5 being *extremely important*) to identify how different criterion (for example, content mastery, effort, improvement, and behaviors) contributed to a student's summative grade. From this survey, the teachers identified that content mastery was the most important criterion when figuring a student's summative grade. However, this was the only criterion when figuring a student's summative grade. However, this was the only criterion that was highly rated that was academically related. The other top four criterion all dealt with student behaviors, including attendance, effort, respect for learning, and respect for authority. The teachers surveyed were asked to identify how summative grades were derived for students, labeling the percentage of the grade that was assigned for academics and for behaviors. This information was used to identify cohort groups of similar grading practices.

Rich (2001) asked the teachers participating in his study to randomly choose 10 students from their first general ability class of the day. The teachers tracked these ten students during the course of the term by using A-B-C-D-F grades for the following criteria: academic achievement, effort, behavior, attendance, and personal values. By doing this, Rich could compare the actual grades attained by students with the teachers' perceptions of their grading practices. Rich looked at the data three different ways. First, he looked at the total cohort of students ($n = 160$). Then he divided the students' data by

their teachers' indication of whether or not 80% or more of their students' grades were based on academic achievement, or if less than 80% of their students' grades were derived from academic achievement. These divisions were made for further data analysis.

In analyzing the data, Rich (2001) noted that for the total cohort group, academic achievement was the highest criterion for factoring student grades. However, after splitting the cohort group into those who identified that academic achievement contributed to greater or less than 80% of the student's summative grade, Rich concluded that those teachers who identified they used academic achievement more than 80% of the time (Group A), in actuality used effort as the biggest criterion for student grades. Conversely, those teachers who identified that academic achievement constituted less than 80% of student's grade (Group B) actually used academic achievement as their highest criterion. In fact, Group B's ranking of the nonachievement criterion were all lower than those in Group A.

As district leaders embark upon an examination of assessment and grading practices in a system, discussions revolving around key issues are critical. As stakeholder groups meet to address issues related to the initiative, valuable insights come to the surface. Through these insights, the steering committee for the initiative will be able to plan meaningful conversations, educational seminars, and develop a timeline for rolling the initiative out. The conversations will likely bring issues to the surface that the steering committee had not thought about when planning.

Use of Averaging or Points

Averaging points, or converting points into percents and then averaging them, are two methods used by teachers to determine a student's summative grade. These are the simplest forms of criterion referenced grading practices, which imply that the student is being assessed against a specific learning goal. Due to the fact that averaging percents involves converting the points attained for each assignment, and converting them to a percentage, the terms will be used interchangeably in this paper.

This grading practice began during World War I when the Army needed a quick and efficient way to assess the competencies of their recruits. The Army used multiple choice assessments that were easily scored giving one point for correct answers, and no points for incorrect answers. The Army developed a range of average scores to place recruits in specific fields. This method proved to be so successful, that multiple choice assessments became a widely used tool in assessment (Marzano, 2000). Over the years, the multiple choice assessments evolved into assessment questions that could be marked as right or wrong. Types of assessments included in this category include true/false, fill in the blank, short essay, or matching.

Research indicated that many teachers use criterion-referenced approaches to grading, which is at the heart of a standards-based reporting system. However, it also contended teachers may not be using the approach in the manner that is most valid to reporting student progress (Busick, 2000; Guskey & Bailey, 2001; Marzano, 2000; O'Connor, 2002). Teachers often designate a grade to every assignment that is given by

awarding points or percentages to the work. This is not the intent of a criterion-referenced grading practice. In its truest form, a criterion-referenced assignment should only cover a very specific learning objective, for example, a standard. If a teacher could isolate assignments that dealt with the same topic, it would then be appropriate to average point or percents on these assignments. Weights could be given to the assignments that fell into a group, thereby allowing the teacher to indicate the amount of material that was covered on that specific topic. When specific assignments dealt with the same topic were averaged, a student's true academic achievement of that topic could be deduced. However, a teacher would need to isolate each of the specific skill sets for that class and report student progress towards each individual skill set. The immediate concern that arises from this situation is how to combine the individual achievement data to form a summative grade that is a true reflection of the student's core knowledge in the subject area.

Research centered on teachers and student perceived benefits of using averaged points or percents has also been conducted. When surveyed, both teachers and students reported that they liked the use of points or percents because it allowed the students to know exactly where they stood in the class (Guskey & Bailey, 2001). Students reported that they liked the system because they knew that all their work factored into their final grade, and that if enough assignments were given in a class, a few poor scores would not impede their attainment of a *good* grade (Busick, 2000). The actual findings of research centered around these perceptions showed that while averaging points and percents were an easy way to figure a student's summative grade, students did not necessarily come out

ahead when actual grades were figured (Busick). Students who started slowly in a class but gained steady progression often were unable to attain a *good* grade due to the fact that they had struggled with the concepts presented at the beginning of the course. The summative grade often times was not a true reflection of the student's actual knowledge attainment of the topic area. O'Connor (2004) stated that the median of weighted percentage grades should be used to give a statistically accurate view of a student's academic achievement. Davis (1993) advocated the use of points in classrooms as a way of letting students know an acceptable range of satisfaction.

Guskey and Bailey (2001) contended that the blanket labels like *gifted* or *learning disabled* are sometimes affixed to students based solely on the points they achieve in a classroom, without an in depth look at the student's knowledge attainment in the academic area. It is also important to remember that while a student may excel or struggle in certain academic areas, this label should not travel with the student across all academic areas without sufficient data to support such a move. Kohn (1999) conducted research that revealed "the use of traditional grades are likely to lead to three separate results: less impressive learning, less interest in learning, and less desire to do challenging learning" (p. 43).

Change Theory

Administrators in school systems often find themselves perceived as soothsayers to many different stakeholders when a new initiative is introduced within the system. When addressing stakeholder groups, administrators need to have both a strong research

base and a vision on how the initiative will better the school system for students. When vision and research are presented in conjunction, discussion should ensue that will look at the initiative through many lenses, thus narrowing the focus and providing the insight to make an informed decision for the betterment of the students in a district.

Currently, an area of contention in education is the differing use of assessment and grading practices to report student achievement. Researchers have spent an inordinate amount of time exploring the concept of grading practices and their intentions, and within their research lay some reflective pieces for different stakeholder groups including parents, staff, students, and administrators. In a plan for instituting a standards-based reporting system within a district, one of the stakeholder groups should be comprised of teachers. Within this group, a significant amount of time spent discussing grading practices and their relationship to a standards-based reporting system will facilitate implementation.

In order to make grading practices more consistent within buildings and school districts, critical discussions must take place revolving around data gathered on grading practices. The first pieces of data that must be collected from teachers is what constitutes an A on a classroom assignment; what measurements are used to base a student's grade; surveying how grades are figured; and finally, if or how the use of zeroes are employed in the classroom. From the gathered data, conversations can begin. The focus of the conversations should revolve around the question, "How does this school share what students know and are able to do those with vested interest?"

Three arguments that commonly arise when the conversation of grading practices takes place in a district revolve around the use of zeroes, grading on a curve, and separating behavior from academics when figuring a student's grade. The research base for these issues should be reviewed, synthesized, and presented with the intention that the stakeholder groups will make an informed decision about grading practices.

When looking at grading practices within a school, an important discussion is how grades are determined within a classroom. Typically, two types of grading practices will prevail; norm referenced scoring and criterion referenced scoring. In criterion referenced scoring, the student's academic achievement is measured against a stated objective or set criteria. In norm referenced assessments, the student's academic achievement is contingent upon the success of other students in the classroom. Many times a bell-shaped curve is the representative of norm-referenced assessments. Guskey (1996) pointed out:

Grading on a curve makes learning a highly competitive activity in which students compete against one another for the few scarce rewards (high grades) distributed by the teacher. Under these conditions, students readily see that helping others become successful threatens their own chances for success. As a result, learning becomes a game of winners and losers; and because the number of rewards is kept arbitrarily small, most students are forced to be losers. (pp. 18-19)

By diminishing the use of norm referenced grading systems, consistency will be a by-product. When a norm referenced system is in place, hypothetically, a student could receive a failing mark in one class with a particular score, while with the exact score, they could receive the best grade in the same class two periods later. The ramifications of

norm referenced scoring for students who speak English as a second language and identified special education students should also be taken into consideration. Finally, if a school district or teacher touts high expectations for all students, can norm referenced assessing truly take place?

Lewin (1947) proposed that organizations be unfrozen, changed, and then refrozen. The key to the change process, whether it was at an individual or group level involved “painfully unlearning without loss of ego identify and difficult relearning as one cognitively attempted to restructure one’s thoughts, perceptions, feelings, and attitudes” (as quoted in Schein, p. 1). Further, Lewin felt there were two ways to effect change within an organization. One involved forcing change on individuals, while holding them accountable for the change, while the other involved “removing the restraining forces that inhibit people from understanding or being successful with the proposed change” (McREL, 2006, p. 26).

Many change theorists agreed that the psychological nature of change needs to be addressed (Bridges & Mitchel, 2000; Fullan, 2002; Heifetz & Linsky, 2004; Rogers, 2003). At the psychological awareness stage of the change theory, data should be collected as to the readiness level of the organization to accept change. Several surveys have been developed to acquire the data (Bay Area School Reform Collaborative [BASRC], 1998; Bridges, 2001; Howard, Howell, & Brainard, 1987) with the intention of collecting data on the organizations’ culture and willingness to adapt to change.

However, in addition to the psychological nature of change, some change theorists placed equal importance on defining the term change for the individual or organization, placed equal importance on defining the term change for the individual or organization, thus making participants more cognitively aware of the expectations. For example, Bridges and Mitchell (2000) advocated that leaders distinguish the difference between transitions and change with their staffs. Bridges (1991) defined transition as “the psychological process that people go through as they experience change,” while change is the actual event that occurs. Cuban (1992, 1996, 1997) based his work on the previous work of Watzlawick, Weakland, and Fisch (1974) and discussed the importance of identifying the change that proposed to occur within an organization as first or second order change. First order change is “the most obvious next steps to take in a school or district” (Marzano et al. 2005, p. 66). These changes do not seem to vary far from the existing norms within an organization and align with individual’s personal beliefs and values. Second order changes, conversely, seem to make an apparent break from past practices, require a new skill or knowledge set to be acquired, and may not align closely with an individual’s personal belief and value system.

Marzano et al. (2005) defined this change as a deep change that “alters the system in a fundamental way, offering a dramatic shift in direction and requiring new ways of thinking and acting” (p. 66). In the Balanced Leadership training offered by McREL, a survey is presented that can be used with individuals within an organization to determine whether a proposed initiative would be considered a first or second order change for the organization. Because each school building has its own distinct culture and climate, there

is not a pervasive method to predict whether or not an initiative will be viewed as first or second order change for the building. This survey is included in Appendix A.

School culture and readiness for change impacts organizations on many levels. For example, Guskey and Sparks (1996) cited the research of Crandall, Eisemann, and Louis (1986) who contended that the greatest student achievement gains are a result of staff development that did not veer far from individual's beliefs and values. Researcher Leithwood (2002) noted that, "staff in especially productive schools typically holds norms of continuous improvement and professional growth as well as norms of mutual respect" (p. 99). These researchers suggested that if organizations were asked to change too many aspects of their culture at once, it would be tempting to return to former habits and not implement new thinking and initiatives within the organization, which takes organizations back to the transitioning state of the change process.

There are several researchers who offered characteristics and traits of effective leaders (Marzano et al., 2005; McEwan, 2003; Reeves, 2006; Schmoker, 2006) and each offered how these traits could be used to foster school improvement initiatives. However at the heart of the previously mentioned research is the climate and culture of the organization when organizations begin to think about a change. Through a thorough analysis of school climate and culture, as well as readiness for change data, instructional leaders can develop an action plan that will best meet the needs of the organization in regards to sustainable school improvement changes.

CHAPTER 3: RESEARCH METHOD

Introduction

In 1990, the National Education Goals and Indicators were announced by the President and state governors to address educational issues including (a) school readiness; (b) school completion; (c) student achievement and citizenship; (d) teacher education and development; (e) math and science; (f) adult literacy and lifelong learning; (g) safe, disciplined, and alcohol and drug free schools; and (h) parental participation. The federal legislators set the year of 2000 as the year that the indicators for each of these goals should be met. States set forth to develop standards and benchmarks for the students in their schools that would indicate academic achievement as well as attainment of the National Education Goals and Indicators. When creating standards and benchmarks, states were also aligning federal programming with curriculum, instruction and assessment that occurred in schools. Iowa, a local control state, left this task to each of the 559 school districts in the state.

One such district is an accredited nonpublic system. This district developed standards, benchmarks, and grade level/course level expectations (GLEs/CLEs) four years ago for all content areas K-12, and has been working towards a standards-based reporting system K-12 through the use of a standards-based reporting committee. During the fall of 2007, all of the kindergarten through fifth grade teachers in the district began piloting a reporting form that was reflective of a standards-based reporting system and its associated assessment and grading practices. It was during the course of this

implementation that teachers discovered there was a disconnect within the system with regard to the use of academic and nonacademic factors when determining students' grades.

O'Connor (2002) identified eight assessment guidelines that often spark essential dialogue when a school district is considering a standards-based reporting system (p.46). Based on previous research conducted by the researcher (Lindahl & Roorda, 2006), the secondary school teaching staff of this nonpublic school system in Iowa is in conflict with O'Connor's guidelines and the current assessment and grading practices in some of their classrooms. These conflicts may result in students receiving grades that may not be a true reflection of their academic achievements.

A study examining secondary school teachers' use of academic and nonacademic factors (dependent variable) when determining students' grades will add to the body of evidence being gathered around these topics for several reasons. First, by analyzing teachers' use of academic and nonacademic factors before and after professional development on assessment and grading practices will provide heightened awareness of the actual components of a final grade to stakeholders within the district. Second, by comparing the perceived components of a grade, the case for implementing a standards-based reporting system may be built. Third, the data collected from this study could be used to expand research on the actual achievement data colleges are analyzing to determine students' admissions to postsecondary educational institutions. Finally, the data collected from this study will provide insight into the purposes for assessment and grading practices in secondary schools in the United States.

There is a problem in assessment and grading practices as they relate to determining summative student grades in the secondary setting. That problem, specifically, is a misalignment between assessment and grading practices and the reporting of academic achievement. Currently, both academic (student achievement) factors and nonacademic (homework, participation, behaviors) contribute to the determination of grades. However, if teachers are reporting academic achievement progress, nonacademic factors should not be reflected in students' grades. This problem impacts secondary school students because their grades may not be a true reflection of their academic achievement. There are many possible factors contributing to this problem, among which are assigning extra credit to students, assessing students with the inclusion of nonacademic factors, and the differences in assessment and grading practices for students that occur within a department and/or school setting. This study will contribute to the body of knowledge needed to address this problem by looking at the assessment and grading practices of a sample of secondary teachers when determining final student summative grades. The study will also contribute to the body of knowledge on whether there are differences in the assessment and grading practices for different academic content areas.

Role of the Researcher

Prior to a new role within the agency for which she is employed, the researcher was a school improvement consultant for the Acme Education Corporation, which serves as an intermediate agency between local school districts and the State Department of Education. The researcher worked with the nonpublic system by conducting professional

development within the system. During the study, the researcher worked with a third party to complete data analysis measures on the survey data collected.

Research and Design Approach

The purpose of this repeated measures quantitative study was to compare secondary teachers' use of academic and nonacademic factors when reporting student achievement before and after professional development on assessment and grading practices in a standards-based reporting system. A quantitative research design was chosen over a qualitative research design because data analyzed were able to be generalized and objective, while a qualitative study dealt more with subjective data that was not able to be generalized. Further, a quasi-experimental repeated measures study was chosen as the best methodological match for this study based upon the fact that the research was conducted using a pre-/post-test design. A descriptive research design was rejected by the researcher because the purpose of the study was not to develop theory or identify problems with current practices of the participants. A correlational research design was rejected because the study was not intended to solely be conducted after the professional development had taken place.

This study included secondary teachers (6th-12th grade) in a nonpublic school system in the Midwest completing the Teacher Survey on Grading Practices, which is a previously validated pre- and posttest survey developed by and used with permission. This survey identified the use of academic and nonacademic factors when determining summative student report card grades. In addition, a validated frequency count survey designed by and used with permission from McREL (2006) provided data on the

magnitude of change for an impending initiative, in this case, a standards-based reporting system with its associated assessment and grading practices.

When using the Teacher Survey on Grading Practices, a Likert scale survey developed and validated by Rich (2001), teachers completed the survey using an online survey vehicle, Test Pilot. For the purpose of this study, academic factors are defined as those indicators that are directly tied to student academic achievement, including subject specific content learned, reasoning and thinking skills demonstrated, and communication skills demonstrated. Nonacademic achievement factors are defined as indicators that could be included in a grade such as behavior, attendance, participation, and work completion. Teachers also answered several questions that led to the magnitude of change associated with the implementation of a standards-based reporting system and its associated assessment and grading practices.

Group A O_1 ——— X ——— O_2

Group A represents the teachers in the study who complete the pretest (O_1), then have additional training (X) on assessment and grading practices as they relate to a standards-based reporting system, and finally take a posttest (O_2). The researcher chose this graphic to represent the study as it reflects the pre/post test format that all participants completed, while also graphically representing the independent variable, professional development, with regard to assessment and grading practices in a standards-based reporting system. The researcher then compared the data gathered about perceptions of the staff around student assessment, the dependent variable, before and

after participation in professional development revolving around assessment and grading practices in a standards-based reporting system.

A threat to validity may be the exposure all teachers have had in regard to the effects of different assessment and grading practices on students' summative grades. In addition, there is no way to measure the truthfulness of the responses provided by respondents. This study specifically addressed the following research questions:

Question 1: With the implementation of a standards-based reporting system and its associated assessment and grading practices, is there a significant difference in the perception of nonpublic secondary teachers about the magnitude of change involved?

The null hypothesis for this question is (H_0): With the implementation of a standards-based reporting system and its associated assessment and grading practices, there is no significant difference in the perception of nonpublic secondary teachers about the magnitude of change involved.

The alternative hypothesis for this question is (H_1): With the implementation of a standards-based reporting system and its associated assessment and grading practices, there is a significant difference in the perception of nonpublic secondary teachers about the magnitude of change involved.

Independent variable - implementation a standards-based reporting system and its associated assessment and grading practices

Dependent variable - perception of nonpublic secondary teachers

Question 1: With the implementation of a standards-based reporting system and its associated assessment and grading practices, is there a significant difference in the perception of nonpublic secondary teachers about the magnitude of change involved?

The null hypothesis for this question is (H_0): With the implementation of a standards-based reporting system and its associated assessment and grading practices, there is no significant difference in the perception of nonpublic secondary teachers about the magnitude of change involved.

The alternative hypothesis for this question is (H_1): With the implementation of a standards-based reporting system and its associated assessment and grading practices, there is a significant difference in the perception of nonpublic secondary teachers about the magnitude of change involved.

Independent variable - implementation a standards-based reporting system and its associated assessment and grading practices

Dependent variable - perception of nonpublic secondary teachers

Question 3: Is there a significant correlation between the teachers' content area taught and their assessment and grading practices?

The null hypothesis for this question is (H_0): There is no significant correlation between the teachers' content area taught and their assessment and grading practices.

The alternative hypothesis for this question is (H_1): There is a significant correlation between the teachers' content area taught and their assessment and grading practices.

Independent variable – teachers' content area taught

Dependent variable – assessment and grading practices

Setting and Sample

The purpose of this quantitative repeated measures study using a pre/posttest design is to compare assessment and grading practices of a random sample of 39 secondary school teachers from three buildings in a Midwestern nonpublic school system (with approximately 6, 278 students in 18 buildings) both before and after professional development revolving around assessment and grading practices (dependent variable) in a standards-based reporting system.

Group A O_1 ——— X ————— O_2

Group A represents the teachers in the study who complete the pretest (O_1), then have additional training (X) on assessment and grading practices as they relate to a standards-based reporting system, and finally take a posttest (O_2). The researcher chose this graphic to represent the study as it reflected the pre- and posttest format that all participants completed, while also graphically representing the independent variable, professional development, in regards to assessment and grading practices in a standards-based reporting system. The researcher compared the data gathered about perceptions of the staff around student assessment and grading practices, the dependent variable, before and after participating in professional development revolving around assessment and grading practices in a standards-based reporting system.

A threat to validity may be the exposure all teachers have had to the effects of different assessment and grading practices on students' summative grades. The nature of the study may have been impacted by the limited number of participants in the study

compared with the number of secondary school teachers across the nation if a national generalization is inferred from the data. In addition, there was no way to measure the truthfulness of the responses provided by respondents.

The simple random sampling for this study was drawn from a list of 39 secondary school teachers employed in three buildings in a Midwestern nonpublic school system. The nonpublic system is comprised of 16 K-8 schools and two high schools that serve student 9-12. The system is accredited by the State Department of Education. All teachers in the system are licensed by the board of educational examiners in the state. During the 2005-2006 school year, the total enrollment for the system was 6, 278. Of those students, 4, 874 were enrolled in the K-8 system, and 1,405 were enrolled in grades 9-12. During that school year, there were 378 full time and 63 part time teachers with a student-teacher ratio of 16 to 1 (Nonpublic School, 2007).

Every secondary teacher from three buildings was identified from a composite list. The staff member's names were then cut into strips with one name on each strip. Names were then put into a box, thoroughly mixed, and then drawn out. The participants selected were then asked to complete an online survey about the weight they gave to academic and nonacademic factors (behaviors, homework completion, participation, absences, etc.), when determining final summative grades for students, as well as their perception of the magnitude of change associated with implementing a standards-based reporting system in the secondary setting. Another factor that was examined in the study included the teachers' content area taught. The total number of participants available in the study was 39, and a sample size of 35 was determined using the sample size

calculator for 5% error and 95% confidence level. The researcher realized that this sample size is small for a quantitative study, but it was appropriate based on the fact that the context of this particular study limited the researcher to a small sample size. In other words, the school system is small, yet it was investigating implementation of a standards-based reporting system in the secondary, which made it appropriate for the study.

Treatment

Professional development centered on assessment and grading practices within the secondary section of the nonpublic district was used as the independent variable in the study. The professional development explored (a) current assessment and grading practices within the secondary system, (b) potential obstacles in implementing the assessment and grading practices agreed upon by the administrators and staff, and (c) present training on the assessment and grading practices as outlined.

The professional development sessions were delivered at two separate sites due to the geographic size of the nonpublic system. Sessions were duplicated on two separate days each time professional development sessions were presented. The sessions ranged from two hour to full day (8 hour) training sessions. During each of the sessions, there were opportunities for training, discussion, and recommendations by staff to take back to administrators regarding the use of assessment and grading practices in a standards-based system.

The professional development sessions included, but were not limited to a discussion of the magnitude of change involved with the implementation of a standards-based reporting system; analysis of current practices around assessment and grading

practices; training on assessment and grading practices based on the work of a variety of researchers (Brookhart, 1993; Busick, 2000; Guskey & Bailey, 2001; Marzano, 2003; Marzano, 2006; O'Connor, 2002; Stiggins, 1997; Tomlinson & McTighe, 2006; Wiggins, 1994; Wormeli, 2006) and horizontal and vertical grade level discussions on the consistent use of assessment and grading practices.

Instrumentation and Materials

A three-part instrument was designed for data collection for this research study. The first part of the instrument was a survey that evaluated the magnitude of change associated with the implementation of a standards-based reporting system. This survey was developed by researchers at McREL and was used with permission. The study was validated through the use statistical analysis and employed criterion validity to predict the magnitude of change associated with impending initiatives. The other two portions of the instrument were adapted and used with permission from a survey instrument used by Rich (2001) to determine the weight given to hidden curricular items such as participation, by teachers. The Likert scaled survey was used to capture teachers' assessment and grading practices involving both academic and nonacademic factors before they began professional development focused on assessment and grading practices in a standards-based reporting system. The survey was completed again after the conclusion of the professional development sessions.

Rich validated the instrument by presenting the survey to multiple secondary teachers in the state of New Jersey and revising it based on feedback collected in interview sessions after the pilot of the survey. The survey itself is published in Rich's

doctoral study, *Hidden Factors in Teachers' Secondary Grading Practices* (2001). Internal and external validity were analyzed based on the participants that Rich in designing and implementing his study. Rich used the test/retest administration for reliability in his instrument. A copy of the magnitude of change survey is found in Appendix A and the adaptation of Rich's survey instrument is found in Appendix B.

Data Collection Procedures

The data that were collected for this study were gathered using a free survey tool available to the district called Test Pilot. Test Pilot, a survey vehicle similar to Survey Monkey, was housed at the Acme Learning Corporation, an intermediate educational unit, and is provided as a free service to districts within Acme's boundaries. Prior to administering the survey to the teaching staff, the researcher loaded Rich's survey and the McREL survey into the Test Pilot program. Prior to the completion of the 2007-2008 school year, the study participants independently took the Test Pilot survey on the internet according to directions provided.

Administration of the online survey instrument took place prior to the completion of the school year on May 30, 2008. Teachers were given a sheet that explained the purpose of the survey, as well as how to access the Test Pilot survey online. The sheet also included frequently asked questions about Test Pilot in order to help with basic troubleshooting.

The researcher worked collaboratively with the administrators of the nonpublic system to collect the data at each of the individual sites. The data was collected using the Test Pilot online survey tool through a partnership with the Acme Learning Corporation.

Prior to the administration of the survey, administrators were shown the survey and given a summary of the data that would be available to them after the survey was completed by secondary school staff both before and after professional development activities revolving around assessment and grading practices. After the initial pretest (before professional development began), the researcher met with the administrators to discuss results and how they might impact action planning for professional development provided to secondary teachers. The researcher met with administrators individually after the second survey was administered to discuss the relationship between assessment and grading practices and professional development on assessment and grading practices in a standards-based reporting system. The researcher will also meet with administrators collectively to discuss the results of the study.

Secondary school teachers completed a sign-in sheet prior to accessing the Test Pilot survey. Since the initial (pretest) survey was completed during the school day, little absenteeism was predicted. However, if a staff member was absent on the day the survey was administered, the administrator(s) in the building provided initial follow up to the absent staff member to complete the survey within three days of their return. If the survey was not completed in the timeframe, the researcher directly contacted the staff member to complete the survey. Once the random sample of 35 teachers was collected, no further follow up for delinquent surveys was conducted.

After both pre-/post-survey was collected, the researcher used the SPSS program to conduct the data analysis portion of the study.

Data Analysis Plan

The first portion of the survey involved a frequency count indicating the number of teachers who perceived the implementation of a standards-based reporting system with associated assessment and grading practices as a first or second order change.

Descriptive statistics including frequencies and means were used to analyze the data collected regarding the magnitude of change involving the standards-based reporting initiative. These descriptive statistics were chosen because the researcher wanted to present the entire distribution in a table mode, which would indicate the use of a frequency distribution table. A bar graph and histogram were rejected due to the fact that the information represented was not in nominal, ordinal, interval or ratio scales (Gravetter & Wallnau, 2005, p. A-50). The researcher chose to summarize the entire distribution using means, which was chosen because the central tendency was reported using an interval scale. Due to the fact that the researcher wanted to compare interval scores, and not the most frequently occurring data points in a region, the mode was not chosen as a descriptive statistic to be analyzed. In addition to frequency counts and means, the researcher also completed the chi-square for independence with 95% significance for Part I of the Standards-Based Grading Practices Survey Teacher Questionnaire: Adapted from Rich's Teacher Survey on Grading Practices and analysis for correlations was done through the use of Spearman's Correlation for the items on Part 2 of the Standards-Based Grading Practices Survey Teacher Questionnaire: Adapted from Rich's Teacher Survey on Grading Practices.

The items on Part I of the Standards-Based Grading Practices Survey Teacher Questionnaire: Adapted from Rich's Teacher Survey on Grading Practices collected data on a Likert scale, which was analyzed as interval data. Since the data collected was nonparametric, a chi-square with 95% significance was the best analytical match. Knowing that "all parametric test place stringent restrictions on the sample data and the populations distributions being considered" (Gravetter & Wallnau, 2005, p. A-53), parametric tests such as single sample t tests and z-score tests were rejected. As the study was a repeated measures study with the same subjects completing the survey twice, the researcher analyzed the difference in scores for the two surveys. A Spearman's Correlation was utilized in the second portion of the survey as the researcher was measuring relationships on an ordinal scale in Part 2 and the responses from Part 1. The researcher rejected using a Pearson Correlation or Regression analysis due to the fact that the scores reported by participants were not numerical values from interval or ratio scales (p. A-58).

Validity of Study

The study was validated through the use statistical analysis and employed criterion validity to predict the use of grading and assessment practices of secondary teachers after participating in professional development centered on grading and assessment practices in a standards-based reporting system. Internal validity was ensured through the use of an online survey system and tools developed to provide consistent administration of the survey tool. External validity was present in the fact that the survey was designed for all secondary teachers, regardless of geographic location or school

system size. A threat to validity was the exposure all teachers had with regard to the effects of different assessment and grading practices on students' summative grades. The nature of the study may be impacted by the limited number of participants in the study compared with the number of secondary school teachers across the nation if a national generalization is inferred from the data. In addition, there was no way to measure the truthfulness of the responses provided by respondents.

Reliability of Study

Rich (2001) utilized test-retest to measure reliability in his study. The researcher in this study also utilized test-retest to measure reliability in the study. The use of the online survey system, Test Pilot, provided technological tools to ensure that data were kept in the same manner each time the survey was taken.

Protection of Participants' Rights

Many measures were taken by the researcher to protect the rights of the participants of the study. First, no data were collected until the study was approved by the Institutional Review Board (IRB) with the approval number of 05-04-08-329541.

After receiving approval to conduct the study, the researcher reviewed the purpose and details of the study at meetings across the system. The researcher advised the secondary teachers that their participation in the study would be entirely voluntary. The secondary staff members were given the researcher's name, email address, and phone number. Staff members who expressed an interest in participating in the study or in learning more about the nature and purpose of the study were asked to contact the researcher by telephone or via email. This ensured that each participant's identify was

not known to his/her colleagues. Each participant was given an information sheet (Appendix C) and was asked to sign an informed consent form prior to participation in the study. Participants were informed of their ability to request a copy of the data collected after the completion of the research.

Confidentiality

The data collection instrument used in the study was an online survey that did not ask for demographic information that might reveal the participant's identity. The participants were assured that their responses, as linked to identity would not be shared with administrators or to fellow colleagues. Participants were informed that the consent forms signed by those involved would be kept for seven years and then destroyed.

Risks and Benefits

This research study asked secondary teachers to identify their use of academic and nonacademic factors in assessment and grading practices prior to attending professional development on these topics. The researcher reminded all participants that they had the freedom to withdraw from the study at any time. There were direct benefits to the secondary teachers involved in this study. Benefits included discussion of current assessment and grading practices found within a building and the system; the ability to provide feedback into the assessment and grading practices that will be used in the secondary setting of the nonpublic system; and the opportunity to network with secondary teachers from other buildings during professional development sessions. The secondary teachers' participation in the research may have gleaned a deeper insight into their own assessment and grading practices.

The sample for this study consisted of secondary teachers in a nonpublic school district. Therefore, the data did not reflect experiences of elementary teachers in a nonpublic district, nor elementary or secondary teachers in public school districts. It is important to keep in mind that the data collected may have been limited by the participants' personal philosophies and willingness to respond. Secondly, the secondary teacher volunteers may represent a special group with attitudes different from those who did not volunteer to participate in the study.

CHAPTER 4:

RESULTS

Questions & Hypotheses

As outlined in Chapter 1, this study was conducted to look at the impact of professional development on standards-based reporting systems and their associated grading and assessment practices in a secondary nonpublic school system. This chapter is organized in terms of the three specific research questions that were presented in Chapter 1:

Question 1: With the implementation of a standards-based reporting system and its associated assessment and grading practices, is there a significant difference in the perception of nonpublic secondary teachers about the magnitude of change involved?

The null hypothesis for this question is (H_0): With the implementation of a standards-based reporting system and its associated assessment and grading practices, there is no significant difference in the perception of nonpublic secondary teachers about the magnitude of change involved.

The alternative hypothesis for this question is (H_1): With the implementation of a standards-based reporting system and its associated assessment and grading practices, there is a significant difference in the perception of nonpublic secondary teachers about the magnitude of change involved.

Independent variable - implementation a standards-based reporting system and its associated assessment and grading practices

Dependent variable - perception of nonpublic secondary teachers

Question 1: With the implementation of a standards-based reporting system and its associated assessment and grading practices, is there a significant difference in the perception of nonpublic secondary teachers about the magnitude of change involved?

The null hypothesis for this question is (H_0): With the implementation of a standards-based reporting system and its associated assessment and grading practices, there is no significant difference in the perception of nonpublic secondary teachers about the magnitude of change involved.

The alternative hypothesis for this question is (H_1): With the implementation of a standards-based reporting system and its associated assessment and grading practices, there is a significant difference in the perception of nonpublic secondary teachers about the magnitude of change involved.

Independent variable - implementation a standards-based reporting system and its associated assessment and grading practices

Dependent variable - perception of nonpublic secondary teachers

Question 3: Is there a significant correlation between the teachers' content area taught and their assessment and grading practices?

The null hypothesis for this question is (H_0): There is no significant correlation between the teachers' content area taught and their assessment and grading practices.

The alternative hypothesis for this question is (H_1): There is a significant correlation between the teachers' content area taught and their assessment and grading practices.

Independent variable – teachers' content area taught

Dependent variable – assessment and grading practices

Data Presentation

Thirty-five teachers (89.7%) responded to this survey. Of this group, 28.6% ($n=10$) were male and 71.4% ($n=25$) were female. Twenty percent ($n=7$) of the participants held a BA/BS as their highest degree; 34.4% ($n=12$) held hours beyond their BA/BS, but not their MA/MS; 20.0% ($n=7$) reported a MA/MS degree; 22.9% ($n=8$) have hours beyond their MA/MS, but do not hold Ed.D or Ph.D degree; and 0.03% ($n=1$) did not respond to this question. Table 1 reports frequencies and percents related to highest degree held by participants of the study.

Table 1
Educational Level of Participants in Study

Variable	<i>n</i>	%
BA/BS	7	20
Hours beyond BA/BS, but not MA/MS	12	34.4
MA/MS	7	20
Hours beyond MA/MS, but not Ed.D or Ph.D	8	22.9
Ed.D	0	0
Ph.D	0	0
Total	35	100

Seventeen percent of the respondents ($n=6$) reported they were Language Arts/Reading teachers; 2.9% ($n=1$) were art teachers; 2.9% ($n=1$) foreign language; 17%

($n=6$) math; 8.6% ($n=3$) music; 8.6% ($n=3$) orchestra/band; 2.9% ($n=1$) physical education; 5.7% ($n=2$) religion; 14.3% ($n=5$) science; 8.6% ($n=3$) social studies; and 8.6% ($n=3$) reported *other*.

The first research question addressed in this study sought to clarify what teachers perceived about the magnitude of change associated with implementing a standards-based reporting system in the middle and high school buildings of a non-public school system. The actual research question was, “Is there a difference in the perception of nonpublic secondary teachers about the magnitude of change involved with implementing a standards-based reporting system and its associated assessment and grading practices?”

The null hypothesis for this question is (H_0): There is no significant difference in the perception of nonpublic secondary teachers about the magnitude of change involved with implementing a standards-based reporting system and its associated assessment and grading practices, the dependent variable.

The alternative hypothesis for this question is (H_1): There is a significant difference in the perception of nonpublic secondary teachers about the magnitude of change involved with implementing a standards-based reporting system and its associated assessment and grading practices, the dependent variable.

As described in Chapter 3, a survey was used to assess the perceptions of the teachers. Teachers were given four questions and asked to provide a “yes” or “no” response. Frequencies of the responses are recorded in Table 2.

Table 2
Frequencies of responses in Determining Magnitude of Change for Implementing a Standards-Based Reporting System in the Secondary Setting (n=35)

Variable	Yes Responses	%	No Responses	%
Would the majority of stakeholders perceive the new knowledge required by the change initiative to be easily learned using existing knowledge and skills	12	34.3	23	65.7
Would the majority of stakeholders perceive the change initiative as an extension of the past?	11	31.4	24	68.6
Would the majority of the stakeholders perceive the change initiative as congruent with their personal values and beliefs?	18	51.4	17	48.6
Would the majority of the stakeholders perceive the change initiative as consistent with the prevailing norms?	19	54.3	16	45.7

In the first two questions, more participants felt the implementation of a standards-based reporting system in the secondary setting (middle school/high school) would not be perceived as a natural extension of existing knowledge or skills (65.7%), nor would it be considered an extension of the past (68.6%). The majority of respondents, however, did report that the implementation of a standards-based reporting system would be congruent with their personal values and beliefs (51.4%) as well as consistent with prevailing norms (54.3%). These data supports the rejection of the null hypothesis for this question.

The second research question sought an understanding of how teachers relate academic and nonacademic factors into a student's summative grade in a class. The data analysis related to the question follows.

Question 1: With the implementation of a standards-based reporting system and its associated assessment and grading practices, is there a significant difference in the perception of nonpublic secondary teachers about the magnitude of change involved?

The null hypothesis for this question is (H_0): With the implementation of a standards-based reporting system and its associated assessment and grading practices, there is no significant difference in the perception of nonpublic secondary teachers about the magnitude of change involved.

The alternative hypothesis for this question is (H_1): With the implementation of a standards-based reporting system and its associated assessment and grading practices, there is a significant difference in the perception of nonpublic secondary teachers about the magnitude of change involved.

Independent variable - implementation a standards-based reporting system and its associated assessment and grading practices

Dependent variable - perception of nonpublic secondary teachers

The instrument used to collect these data was a 16 question online test pilot survey administered prior to professional development centered on standards-based reporting systems and their associated grading and assessment practices, and again at the conclusion of the professional development series. The survey involved teachers ranking different criteria, both academic and nonacademic, in comparison to the amount of

consideration given to the stated objective when determining a student's summative report card grade.

The following five-point Likert scale was used

1 = *Not considered*

2 = *Little consideration*

3 = *Some consideration*

4 = *Moderate consideration*

5 = *Substantial consideration*

Frequencies of responses to the assessment practices in the survey were tabulated and the means and standards deviations are presented in Table 3.

Table 3
Descriptive statistics for comparing pre/post survey data on academic and non academic factors used when determining student summative grades (n=35)

Variable	<i>M</i> Pre	<i>M</i> Post	<i>SD</i> Pre	<i>SD</i> Post
Comparison	2.0286	1.8857	1.09774	1.20712
Mastery	4.2000	4.5714	1.27879	0.94824
Improvement	3.8857	4.4857	1.15737	0.70174
Effort	3.4857	3.1714	1.35845	1.33913
Behavior	2.2571	1.6857	1.29121	1.15737
Absences	2.6000	2.4000	1.28795	1.21752
Tardy	1.9429	1.7714	0.96841	0.97274
Participation	3.1143	1.7714	1.34539	0.97274
Homework	3.0000	2.4286	1.51463	1.33473
Extra Credit	3.1543	3.2286	1.33662	1.16533
Challenge	2.8000	1.5429	1.38903	0.81684
Weight	2.2857	1.6000	1.80801	1.06274

Table 3 indicates that the most important criteria that teachers used when determining summative grades before the professional development opportunities were provided was student mastery ($M = 4.2000$), improvement ($M = 3.8857$), effort ($M = 3.4875$), and the participation of the student in class ($M = 3.1143$). After professional development was completed, teachers indicated that student mastery ($M = 4.5714$) and the amount of overall improvement made by students ($M = 4.4857$) continued to be the top criteria that were considered the most when determining a summative grade. It is

noted that these data indicate the academic attainment related criteria of improvement (pre $M = 3.8857$ /post $M = 4.4857$), and mastery (pre $M = 4.2000$ /post $M = 4.5714$) increased from the beginning of the study to the end. It is also noted that all nonacademic factors on the survey decreased in the amount of consideration teachers gave them when determining a summative grade for students.

A chi-square for independent samples with 95% confidence was also completed for this section of the survey. In comparing these data, there was evidence that a significant difference between the gender of the survey participant and the consideration given to the role of homework in the pre survey with $\chi^2 (4, N = 35) = 12.833, p = 0.012$. When performing a chi-square for independent samples with regard to the survey participants' teaching areas, significant differences were noted for weight in the pre survey with $\chi^2 (40, N = 35) = 80.699, p = 0.039$; tardiness in the post survey with $\chi^2 (40, N = 35) = 44.882, p = 0.040$; and participation in the post survey with $\chi^2 (40, N = 35) = 44.882, p = 0.040$. Finally, in completing a chi-square for independence with 95% confidence in the level of education attained by the study participants and the academic and nonacademic factors considered when assessing and grading secondary students, a significant difference was found in the pre survey around behaviors $\chi^2 (16, N = 35) = 35.458, p = 0.003$ and tardiness $\chi^2 (16, N = 35) = 23.260, p = 0.026$. The post survey indicated a significant difference between educational attainment of study participants and the use of extra credit as a consideration in grading and assessment practices with $\chi^2 (16, N = 35) = 25.734, p = 0.050$.

The third section of the survey related to the gender of the survey participants, their gender, the highest level of education achieved by the participants, and finally, the content area that the participant taught. This information was used to answer the following question:

Question 3: Is there a significant correlation between the teachers' content area taught and their assessment and grading practices?

The null hypothesis for this question is (H_0): There is no significant correlation between the teachers' content area taught and their assessment and grading practices.

The alternative hypothesis for this question is (H_1): There is a significant correlation between the teachers' content area taught and their assessment and grading practices.

Independent variable – teachers' content area taught

Dependent variable – assessment and grading practices

A Spearman's rho correlation analysis was utilized in the second portion of the survey as the researcher will be measuring relationships on an ordinal scale in Part 2 and the responses from Part 1. The researcher rejected using a Pearson Correlation or a Regression analysis due to the fact that the scores reported by participants were not numerical values from interval or ratio scales (Gravetter & Wallnau, 2005, p. A-58).

A significant correlation was determined to exist between the survey participants' teaching area and how tardiness was factored into grading and assessment factors with a Spearman's rho of 0.348 with a $p = .041$. A Spearman's rho of 0.335 with a $p = .049$ was also noted on the pre survey with regard to the consideration given to student

improvement when considering grading and assessment practices. The results led the researcher to reject the null hypothesis.

Interpretation

At the onset of the study, the researcher wanted to investigate the perceived magnitude of change for the implementation of a standards-based reporting system and its associated grading and assessment practices within a secondary non public school system in the Midwest. The researcher solicited and received permission from McREL (2006) to use their Magnitude of Change survey to solicit results from 35 participants. The results of the survey that were presented in this chapter affirm that the implementation of a standards-based reporting system and its associated grading and assessment practices would be considered a second order change for the system. A second order change is noted as being a change that is inconsistent with prevailing norms. Marzano, Waters, and McNulty (2005) defined this change as a deep change that “alters the system in a fundamental way, offering a dramatic shift in direction and requiring new ways of thinking and acting.”

The second question addressed by the researcher dealt with the difference in the assessment and grading practices (dependent variable) of secondary teachers with regards to academic and nonacademic factors when determining students’ grades after receiving professional development (independent variable) about assessment and grading practices in a standards-based reporting system. To this end, an online survey format was used before and after the study participants received professional development on standards-

based reporting systems and their associated grading and assessment practices. The data analyzed indicated rejected the null hypothesis that there would be no difference in the grading and assessment practices and use of nonacademic and academic factors when determining summative student grades after professional development. While mastery of material presented and student improvement were the top criteria considered when determining summative student grades in both the pre- and postsurveys, it was noted by the researcher that data indicated the means for consideration of each of these criteria increased from the beginning of the study to the end. It was also noted that all non-academic factors on the survey decreased in the amount of consideration teachers gave them when determining a summative grade for students.

A chi-square for independent samples with 95% confidence was also completed for this section of the survey. In comparing the data, there were evidence that a significant correlation between the gender, the area of teaching concentration, and highest level of educational attainment of the survey participant and the role of homework, consideration of nonacademic factors, tardiness, and participation when considering student summative grades in either the pre- or postsurvey results. The results of these data rejected the null hypothesis presented.

Finally, the researcher chose to examine if there were any correlations between the area of academic concentration of the survey participants that the variables that were examined in the study. A significant correlation was determined to exist between the survey participants' teaching area and how tardiness and improvement were factored into

grading and assessment factors in a standards-based system. These results led the researcher to reject the null hypothesis.

SECTION 5:

SUMMARY, CONCLUSION, AND RECOMMENDATIONS

Overview

Grading and assessment practices in classrooms have been researched many times during the last decade (Busick, 2000; Marzano, 2000,2006; O'Connor, 2002, 2004; Rich, 2001; Trumbull 2000). In addition, studies around the use of standards-based reporting systems (Carr and Farr, 2000; Flynn, Mesibov, Vermette, & Smith, 2004; O'Connor, 2002, 2004; O'Shea, 2005) have sparked conversations between educators. Embedded in the philosophical discussions about grading and assessment practices, there are also conversations that center around the use of standards as a criterion rather than norm based way of assessing student achievement. There are many viewpoints about the necessity for the use of grades in education, as well as some time-honored grading and assessment practices, such as the use of zeroes or including homework scores as part of a final grade in a course. This appears to be true in the middle and high school arenas of education. This misalignment is also apparent in the nonpublic secondary school in the Midwest, which was the focus of this study.

During the course of the study, the researcher utilized an online testing survey instrument called Test Pilot to collect data on how secondary teachers' perceived the magnitude of change associated with the implementation of a standards-based reporting system in the secondary setting (dependent variable). The survey also included questions on the use of assessment and grading practices before and after professional development sessions (independent variable) were provided on assessment and grading practices

within a standards-based reporting system. This study addressed three primary research questions:

Question 1: With the implementation of a standards-based reporting system and its associated assessment and grading practices, is there a significant difference in the perception of nonpublic secondary teachers about the magnitude of change involved?

The null hypothesis for this question is (H_0): With the implementation of a standards-based reporting system and its associated assessment and grading practices, there is no significant difference in the perception of nonpublic secondary teachers about the magnitude of change involved.

The alternative hypothesis for this question is (H_1): With the implementation of a standards-based reporting system and its associated assessment and grading practices, there is a significant difference in the perception of nonpublic secondary teachers about the magnitude of change involved.

Independent variable - implementation a standards-based reporting system and its associated assessment and grading practices

Dependent variable - perception of nonpublic secondary teachers

Question 1: With the implementation of a standards-based reporting system and its associated assessment and grading practices, is there a significant difference in the perception of nonpublic secondary teachers about the magnitude of change involved?

The null hypothesis for this question is (H_0): With the implementation of a standards-based reporting system and its associated assessment and grading practices,

there is no significant difference in the perception of nonpublic secondary teachers about the magnitude of change involved.

The alternative hypothesis for this question is (H_1): With the implementation of a standards-based reporting system and its associated assessment and grading practices, there is a significant difference in the perception of nonpublic secondary teachers about the magnitude of change involved.

Independent variable - implementation a standards-based reporting system and its associated assessment and grading practices

Dependent variable - perception of nonpublic secondary teachers

Question 3: Is there a significant correlation between the teachers' content area taught and their assessment and grading practices?

The null hypothesis for this question is (H_0): There is no significant correlation between the teachers' content area taught and their assessment and grading practices.

The alternative hypothesis for this question is (H_1): There is a significant correlation between the teachers' content area taught and their assessment and grading practices.

Independent variable – teachers' content area taught

Dependent variable – assessment and grading practices

Collectively, the three questions that were studied dealt with the whether there would be an impact in the use of grading and assessment practices after secondary teachers attended professional development revolving around the implementation of a standards-based reporting system and its associated grading and assessment practices. As

the results of Chapter 4 indicate, for each of the questions researched, the null hypothesis was rejected based on the data collected for the study.

Interpretation of Findings

The first question that the researcher asked of the survey participants centered around the Magnitude of Change that was associated with the implementation of a standards-based reporting system and its associated grading and assessment practices within the secondary unit of a nonpublic school system. This question was posed to substantiate or refute the research completed by Lewin (1947), Cuban (1992, 1996, 1997), and Marzano et al. (2005). Each of these researchers broadened the definitions of incremental and fundamental changes within systems. Those changes that seemed to be a natural extension of the past were deemed to be first order changes (Marzano et al.) while the authors coined the term second order change for those changes that have implications that are very divergent from past practices. The researchers at McREL gave the researcher permission to use their Magnitude of Change Survey (2006). The purpose of the survey is to determine the magnitude of change involved with the implementation of a standards-based reporting system in the secondary unit of a nonpublic school system. The results from chapter 4 indicated that more participants felt the implementation of a standards-based reporting system in the secondary would not be perceived as a natural extension of existing knowledge or skills (65.7%), nor would it be considered an extension of the past (68.6%). The majority of respondents, however, did report that the implementation of a standards-based reporting system would be congruent with their personal values and beliefs (51.4%) as well as consistent with prevailing norms (54.3%).

Based on this information, the null hypothesis was rejected, thus indicating that for secondary staff, the implementation of the standards-based reporting system and associated grading and assessment practices would be termed a second order change, as defined by Marzano et al. (2006). The results also indicated that the concepts and theory behind standards-based reporting systems are congruent with most of the survey participants' philosophical views on teaching and learning, but their perception is that for the system, this would be a change in the standard procedures for grading and assessing. This information, coupled with the work of Guskey and Sparks (1996) who cited the research of Crandall et al. (1986) in their contention that the greatest student achievement gains are a result of staff development that did not veer far from individual's beliefs and values, might implicate the change to the new system may take time, but would be supported by research. The key to leading the effective implementation of the standards-based reporting system is held by the building leadership of each school.

There are several researchers who offered characteristics and traits of effective leaders (Marzano et al., 2005; McEwan, 2003; Reeves, 2006; Schmoker, 2006) and each offered how these traits could be used to foster school improvement initiatives. However, at the heart of the research is the climate and culture of the organization when organizations begin to think about a change. Analysis of school climate and culture, as well as readiness for change data, could assist instructional leaders in developing an action plan to best meet the needs of the organization in regards to sustainable school improvement changes.

The second portion of the survey dealt with the use of academic and nonacademic

factors used when determining summative student grades. Researchers (Busick, 2000; Guskey & Bailey, 2001; Marzano, 2000; O'Connor, 2002) have shown that teachers use nonacademic factors like homework completion, participation, and attendance when determining grades in a greater percentage than academic factors like meeting standards and benchmarks, gains in learning, and criterion-based assessment.

In both the pre- and postsurvey, mastery of material presented and student improvement were the top criteria considered when determining summative student grades the data indicated the means for consideration of each of these criteria increased from the beginning of the study to the end. It was also noted that all non-academic factors on the survey decreased in the amount of consideration teachers gave them when determining a summative grade for students. The data presented would support the research completed by Messick (1989) and Brookhart (1993).

The theory of validity developed by Messick suggested that when “test scores are interpreted and used” (p.5) they should be analyzed for two aspects of validity: the intended function of the score (interpretation or use) and the source of justification (empirical evidence or social consequences). Messick took the two sources of validity, interpretation and empirical evidence, and crossed them in a matrix that has four representative quadrants: construct validity, relevance and utility, value and implications, and social consequences. In his theory, Messick contended that validity is “a unified concept based on the concept of construct validity” (p. 8). Brookhart (1993) has made connections between Messick’s theory of validity and grading practices used by teachers,

noting that student grade's should be reflective of the academic progress attained by students, and not the nonacademic factors that often contribute to grades.

The final question addressed during the course of this study was seeking to determine if there was a correlation between the academic area taught by the survey participant and their responses to the survey. A significant correlation was determined to exist between the survey participants' teaching area and how tardiness was factored into grading and assessment factors with a Spearman's rho of 0.348 with a $p = .041$. A Spearman's rho of 0.335 with a $p = .049$ was also noted on the pre survey with regard to the consideration given to student improvement when considering grading and assessment practices. The results led the researcher to reject the null hypothesis.

Implications for Social Change

A study examining secondary school teachers' use of academic and nonacademic factors (dependent variable) when determining student grades will benefit the body of evidence being gathered around assessment and grading practices in the secondary setting for several reasons. First, by analyzing the use of academic and nonacademic factors when determining students' grades before and after professional development (independent variable) on assessment and grading practices, a heightened awareness of the actual components of a final grade will be available to stakeholders within the system. Based on information gathered in this study, cognitive dissonance was manifested by survey participants. While teachers perceive the implementation of a standards-based reporting system and its associated grading and assessment practices as in line with their personal philosophies, the implementation would go against the status quo of the system

in which they work. Second, by comparing the perceived components of a grade, the case for implementing a standards-based reporting system and its associated assessment and grading practices can be built. Third, the data collected from this study could be used to expand research on the actual achievement data colleges are analyzing to determine students' admissions to postsecondary educational institutions. Finally, the data collected from this study will provide insight into the purposes for assessment and grading in secondary schools in the United States.

There are many possible factors contributing to this problem, among which are assigning extra credit to students, assessing and grading students with the inclusion of nonacademic factors, and the differences in assessment and grading practices for students that occur within a department and/or school setting. This study will contribute to the body of knowledge needed to address this problem by looking at the assessment and grading practices of a sample of secondary school teachers when determining students' grades. The study will also contribute to the body of knowledge analyzing differences in the assessment and grading practices for different content areas. Through this study, and others like it, the educational system of the United States could be impacted.

Recommendations for Action

The data that were collected can be disaggregated in many different fashions to give the administrators multiple lenses to view the data. This information can help drive the movement towards implementation of a standards-based reporting system and its associated grading and assessment practices in several ways. First, administration could use these data to begin a study of leadership styles that are effective when encountering

second order change within systems. Through the completion of a self-study of leadership styles, the administrators of the system could work together to initiate the factors that are associated with successful change initiatives (Marzano et al., 2005; McEwan, 2003; Schmoker, 2006; Reeves, 2006). Second, data collected from the study could be used to begin to differentiate ongoing professional development activities for the secondary teachers. Within the study, several academic and nonacademic factors were identified that are used when determining student grades. Each of these factors could be broken down further and studied by both administrators and teachers in greater detail with the intention of a system-wide grading and assessment philosophy to emerge. Finally, teachers could be broken into departmental groups to discuss how standards-based reporting systems and their associated grading and assessment practices relate to their discipline area.

From the work outlined above, administrators and teachers can work with the already established Grading and Reporting Committee to develop standards of practice for secondary teachers within the system to use when assessing and grading student work according to the standards.

These data sets could also be shared with the local Board of Education and parent groups as a way of sharing needs assessment data for determining professional development opportunities for the administrative and teaching staff within the system and well as for the development for the standards of practice for grading and assessment within the nonpublic system.

Recommendations for Further Study

The results of this study have led the researcher to several other topics that could add to the literature about grading and assessment practices in a standards-based reporting system. First of all, this study could be extended and the participants of the study could track actual grading and assessment practices used with their students. Researcher Rich (2001) had participants in his study track 10 students in one of their preps to see how they used nonacademic and academic factors when determining student's summative grades. Replication of this research could further validate Rich's findings.

Secondly, research around the feasibility and acceptability of the use of an assessment system that did not use grades could identify how post-secondary systems view grading and assessment practices of their high school colleagues. Researcher Kohn (1993) has researched how students move from a standards-based reporting system in high school to a post-secondary setting. A study centered around scholarship attainment of students from a standards-based high school could support the use of standards-based reporting systems, or further the conversations of post-secondary institutions on how students from standards-based reporting systems plug into their academic institutions.

Finally, a qualitative, or mixed-methods study centered on the emotions tied to grades and assessment in secondary institutions could be correlated with the research completed in this study. The psychological effects of grading and assessment on both students and teachers could be studied with the information being used to determine the best way to record student achievement in a secondary setting.

Conclusion

As outlined in Chapter 1, the results of this study could direct future self-reflection and systematic work in the area of standards-based reporting systems and their associated grading and assessment practices. Through self-reflection of practice and a detailed look at the systematic way grading and assessment are completed, meaningful conversations about student achievement may be attained. These conversations could branch into further conversations with parents and community partners. Legislators, who are in charge of setting educational policy could use this research as a springboard for further research across the areas they are policymakers for to make recommendations on consistent use standards-based reporting and associated grading and assessment practices.

The heart of this work lies in the self-reflection and conversations that need to take place about what a grade means to individual teachers, departments, buildings, school systems, and states. State standards are present in all 50 states, but the grading and assessment practices within those states are not consistent. The grading and assessment practices that are associated with these standards and benchmarks warrant further conversation. It is through these powerful conversations that student achievement can be impacted.

REFERENCES

- Agnew, E.J. (1985, April). *The grading policies and practices of high school teachers*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Airasian, P.W. (1994). *Classroom assessment* (2nd ed.). New York: McGraw-Hill.
- Anderson, L.W., Krathwohl, D.R., Airasian, P.W., Cruikshank, K.A., Mayer, R.E., Pintrich, P.R., Raths, J., & Wittrock, M.C. (Ed.). (2001). *A taxonomy for learning teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. New York: Longman.
- Bay Area School Reform Collaborative. (1998). *BASRC Teacher Survey*. Survey. San Francisco, CA. Retrieved January 16, 2007, from http://www.stanford.edu/group/CRC/publications_files/BASRC_Teacher_Scales_2001.pdf
- Bridges, W. (2001). *Assessing transition readiness*. Survey. Retrieved January 14, 2007 from http://www.wmbridges.com/resources/article_way_through.html
- Bridges, W., & Mitchell, S. (2000). Leading transition: A new model for change. *Leader to Leader*, 16, 1-8. Retrieved January 14, 2007, from <http://pfd.org/leaderbooks/L2L/srping2000/bridges.html>
- Brookhart, S.M. (1993). Teachers' grading practices: Meaning and values. *Journal of Educational Measurement*, 32(2), 123-142.
- Busick, K. (2000). Grading and standards-based assessment. In *grading and reporting student progress in an age of standards*, p. 71-86, edited by E. Trumbull and B. Farr. Norwood, MA: Christopher Gordon.
- Canady, R.L., & Hotchkiss, P.R. (1989). It's a good score! Just a bad grade. *Phi Delta Kappan*, 71(1), 68-71.
- Carr, J. and Artman, E.M. (2002). *The bottom-up simple approach to school accountability and improvement*. Norwood, MA: Christopher-Gordon Publishers.
- Carr, J. and Farr, B. (2000). Taking steps toward standards-based report cards. In *Grading and reporting student progress in an age of standards* (pp. 185-208). Norwood, MA: Christopher-Gordon Publishers.

- Carr, J.F. and Harris, D.E. (2001). *Succeeding with standards: Linking curriculum, assessment, and action planning*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Chappuis, S., Stiggins, R., Arter, J.A., & Chappuis, J. (2004). *Assessment for learning: An action guide for school leaders*. Portland, Oregon: Assessment Training Institute.
- Clark, D. (2000, January). Kurt Lewin. *BNET Research Center*. Retrieved January 20, 2006 from nwlinc.com/~donclark/hrd/history/lewin.html
- Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, F., Mood, A. M., Weinfeld, F. D., et al. (1966). *Equality of educational opportunity*. Washington, DC: U.S. Government Printing Office.
- Cuban, L. (1992). What happens to reforms that last? The case of the junior high school. *American Educational Research Journal*, 29(2), 227 – 251.
- Cuban, L. (1996). Myths about changing schools and the case of special education. *Remedial and Special Education*, 17,75-82.
- Cuban, L. (1997). How schools change reforms: The case of Stanford University School of Medicine, 1908-1990. *American Educational Research Journal*, 34(1), 83 - 112.
- Davis, B.G. (1993). *Tools for teaching*. San Francisco, CA: Jossey-Bass.
- Des Moines Diocese (2007). *Facts about our schools*. Retrieved December 6, 2007, from <http://www.dmdiocese.org/Index.aspx?menuitemid=1140>
- DuFour, R., DuFour, R., Eaker, R., & Karhanek, G. (2004). *Whatever it takes: How professional learning communities respond when kids don't learn*. Bloomington: National Educational Service.
- Elliott, D.C., (2005). *Teaching on target: Models, strategies, and methods that work*. Thousand Oaks, CA: Corwin Press.
- Flynn, P., Mesibov, D., Vermette, P.J., & Smith, R.M. (2004). *Applying standards-based constructivism: A two-step guide for motivating middle and high school students*. Larchmont, NY: Eye on Education.
- Fullan, M. (2002). The change leader. *Educational Leadership*, 59(8), 16-20.

- Gravetter, F.J., & Wallnau, L.B. (2005). *Essentials of statistics for the behavioral sciences 5th ed.* Belmont, CA: Wadsworth/Thompson Learning.
- Guskey, T.R. (1996). Reporting on student learning: Lessons from the past – prescriptions for the future. In *Communicating student learning: The ASCD yearbook 1996*, edited by T.R. Guskey. Alexandria, VA: Association for Supervision and Curriculum Development.
- Guskey, T.R. (2000, December). Grading policies that work against standards ... and how to fix them. *NASSP Bulletin*. Retrieved January 21, 2006 from www.looksmartwomensports.com/p/articles/mi_qa3696/is_200012/ai_n8921332
- Guskey, T. R. & Bailey J.M. (2001). *Developing grading and reporting systems for student learning*. Thousand Oaks, CA: Corwin Press.
- Guskey, T.R., & Sparks, D. (1996). Exploring the relationship between staff Development and improvements in student learning. *Journal of Staff Development*, 17(4), 34-38.
- Hart, G. (1996). Grades: Both a cause and result of fear. *Middle School Journal*. March, 59-60.
- Heifetz, R.A., & Linksy, M. (2004). When leadership spells danger. *Educational Leadership*, 61(7), 33-37.
- Howard, E., Howell, B., & Brainard, E. (1987). *Handbook for conducting school climate improvement projects*. Bloomington, IN: The Phi Delta Kappa Educational Foundation.
- Johnson, J.K. (2001). The grading of elementary student performance on a standards-based report card (Doctoral Dissertation, University of Washington, 2001). Retrieved March 14, 2006, from ProQuest Information and Learning database.
- Kohn, A. (1993). *Punished by rewards: The trouble with gold stars, incentive plans, A's, praise, and other bribes*. Boston: Houghton Mifflin.
- Kohn, A. (1999). *The schools our children deserve: Moving beyond traditional classrooms and "tougher standards."* Boston: Houghton Mifflin.
- Lewin, K. (1947). Frontiers in group dynamics: Channels of group life; social planning; and action research. *Human Relations*, 1, 143-153.

- Liethwood, K. (2002). Organizational conditions to support teaching and learning. In *The keys to effective schools: Educational reform as continuous improvement*, edited by Willis D. Hawley. Thousand Oaks, CA: Corwin Press, Inc.
- Lindahl, R.M., & Roorda, N.L. (2006). *Survey for Des Moines Diocese administrators and teachers*. Survey.
- Marzano, R.J. (2000). *Transforming classroom grading*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Marzano, R.J. (2003). The changing nature of school reform. In *What works in schools: Translating research into action* (pp. 157-158). Alexandria, VA: Association for Supervision and Curriculum Development.
- Marzano, R.J. (2006). *Classroom assessment and grading that works*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Marzano, R.J., Waters, T., & McNulty, B.A. (2005). *School leadership that works: From research to results*. Alexandria, VA: Association for Supervision and Curriculum Development.
- McEwan, E.K., (2003). *Ten traits of highly effective principals: From good to great performance*. Thousand Oaks, CA: Corwin Press, Inc.
- Messick, S. (1989). Meaning and values in test validation: The science and ethics of assessment. *Educational Researcher*, 18(2), 5-11
- Mid-Continent Research for Education and Learning. (2006). *Balanced leadership: School leadership that works: Managing change*. Denver, CO: Author.
- National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform*. Washington, DC: Government Printing Office.
- National Education Goals Panel. (1991). *The national education goals report: Building a nation of learners*. Washington, DC: Author.
- No Child Left Behind Act, 20 U.S.C. § 107-110 (2002).
- O'Connor, K. (2002). *How to grade for learning: Linking grades to standards* (2nd ed.). Thousand Oaks, CA: Corwin.
- O'Connor, K. (2004). Guidelines for grading which support learning and student success. *Classroom Connections International Online Journal*, 3. Retrieved

March 9, 2006, from <http://www.connect2learning.com/pdf/articles/guidelinesForGrading.pdf>

- Olson, L. (1995, June 14). Cards on the table. *Education Week*, 23-28.
- O'Shea, M.R. (2005). *From standards to success*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Reeves, D.B. (2002). *Holistic accountability: Serving students, schools, and community*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Reeves, D.B. (2004). *Accountability for learning: How teachers and school leaders can take charge*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Reeves, D.B. (2004). *101 more questions and answers about standards, assessment, and accountability*. Englewood, CO: Advanced Learning Press.
- Reeves, D.B. (2006). *The learning leader: How to focus school improvement for better results*. Thousand Oaks, CA: Corwin Press.
- Rich, R.H. (2001). Hidden factors in teachers' secondary grading practices (Doctoral Dissertation, Seton Hall University, 2001, UMI 3036925). Retrieved March 14, 2006, from ProQuest Information and Learning database.
- Rogers, E.M. (2003). *Diffusion of innovations*, 5th ed. New York: Free Press.
- Schein, E.H. (n.d.). *Kurt Lewin's change theory in the field and in the classroom: Notes toward a model of managed learning*. Retrieved January 14, 2007, from http://www.a2zpsychology.com/articles/kurt_lewin's_chage_theory.htm
- Schmoker, M.J. (2006). *Results now: How we can achieve unprecedented improvements in teaching and learning*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Squires, D.A. (2005). *Aligning and balancing the standards-based curriculum*. Thousand Oaks, CA: Corwin Press.
- Stiggins, R.J. (1997). *Student-centered classroom assessment* (2nd ed.). Upper Saddle River, NJ: Prentice-Hall.
- Tomlinson, C.A., & McTighe, J. (2006). *Integrating differentiated instruction and understanding by design*. Alexandria, VA: Association for Supervision and Curriculum Development

- Trumbull, E. (2000). Why do we grade – and should we? In *grading and reporting student progress in an age of standards*, p. 23-44, edited by E. Trumbull and B. Farr. Norwood, MA: Christopher-Gordon Publishers.
- Waters, J.T., Marzano, R.J., & McNulty, B. (2004a). Developing the science of educational leadership. *Spectrum: Journal of Research and Information*, 22(1), 4-13.
- Waters, J.T., Marzano, R.J., & McNulty, B. (2004b). Leadership that sparks learning. *Educational Leadership*, 81(7), 48-51.
- Wiggins, G. (1994). Toward better report cards. *Educational Leadership*, 52(2), 28-37.
- Wormeli, R. (2006). *Fair isn't always equal: Assessing and grading in the differentiated classroom*. Portland, ME: Stenhouse Publishers.
- Zemelman, S., Daniels, H., & Hyde, A. (1998). *Best practice: New standards for teaching and learning in America's schools*. Portsmouth, NH: Heinemann.

APPENDIX A:

ESTIMATING THE MAGNITUDE OF A CHANGE

Used with Permission from Mid-Continent Research for Education and Learning

Stakeholder Group:

Questions to Determine Magnitude of Change			
What new knowledge is required for your change initiative?	In what ways does the change initiative build on past practices?	In what ways is the change initiative congruent with the stakeholders' personal beliefs?	In what ways is the change initiative congruent with the prevailing norms?
<p>Would the majority of stakeholders perceive the new knowledge required by the change initiative to be easily learned using existing knowledge and skills?</p> <p>___ Yes ___ No</p>	<p>Would the majority of stakeholders perceive the change initiative as an extension of the past?</p> <p>___ Yes ___ No</p>	<p>Would the majority of the stakeholders perceive the change initiative as congruent with their personal values and beliefs?</p> <p>___ Yes ___ No</p>	<p>Would the majority of the stakeholders perceive the change initiative as consistent with the prevailing norms?</p> <p>___ Yes ___ No</p>

If you answered no to any of the questions, your change initiative has second-order implications for the selected stakeholder group

APPENDIX B:

Standards-Based Grading Practices Survey
Teacher Questionnaire

Adapted with permission from Rich's Teacher Survey on Grading Practices

Part I:

Directions: Teachers consider many factors when determining a student's final summative report card grade. Using the following scale, please indicate the importance to you for each of the following descriptors. This survey will be taken before starting professional development activities on standards-based reporting systems, and again at the completion of the professional development activity.

1 = This descriptor is not considered when determining a student's summative grade

2 = Little consideration is given to this descriptor when determining a student's summative grade

3 = Some consideration is given to this descriptor when determining a student's summative grade

4 = Moderate consideration is given to this descriptor when determining a student's summative grade

5 = Substantial consideration is given to this descriptor when determining a student's summative grade

A. How does this student's work compare to the work of others in the class?

1 2 3 4 5

B. Has the student mastered the Class Level Expectations (CLEs) for the course?

1 2 3 4 5

C. Has the student improved during the quarter?

1 2 3 4 5

D. How much effort has the student put forth during the quarter?

1 2 3 4 5

E. How does this student behave in class?

1 2 3 4 5

F. Has the student missed a lot of class?

1 2 3 4 5

G. Is the student tardy to class often?

1 2 3 4 5

H. Does this student participate in classroom discussions?

1 2 3 4 5

I. Does this student turn in homework on time?

1 2 3 4 5

J. Does this student challenge himself/herself to increase their own level of learning in the classroom?

1 2 3 4 5

Part II:

Directions: Please answer the following questions about yourself. The responses from these questions will be used by the researcher to help further quantify the study.

K. Please indicate from the list below, the amount of weight given to non-academic factors (work completion, attendance, participation, extra credit, etc.) in determining students' summative (report card) grade:

1. 0 - 9%
2. 10 - 20%
3. 21 - 40%
4. 41 - 60%
5. 61 - 80%
6. 81 - 100%

L. Please indicate your gender:

1. Male
2. Female

M. Please indicate the highest level of education you have received:

1. BA/BS
2. Hours beyond BA, but not MA/MS
3. MA/MS
4. Hours beyond MA/MS but not Ed.D or Ph.D
5. Ed.D
6. Ph.D

N. Please indicate your primary area of instruction from the list below:

1. Science
2. Social Studies
3. Math
4. Language Arts (Reading)
5. Religion
6. Art
7. Music

8. Vocal
9. PE
10. Orchestra/Band
11. Foreign Language
12. Other

Thank you for your participation!

Adapted from Rich, R.H. (2001). Hidden factors in teachers' secondary grading practices. ProQuest Information and Learning Company. (UMI No. 3036925).

APPENDIX C

Information Sheet

Dear Secondary Teacher:

I would like to invite you to take part in a research study on *The Impact of Professional Development on Assessment and Grading Practices for Secondary Teachers*, which is being conducted as part of partial fulfillment of the requirements for the degree of Doctor of Education through Walden University. The primary researcher is Nicole (Nikki) Lynn Roorda, who is a doctoral student at Walden University

The purpose of this study is to determine if there is a difference in the use of assessment and grading practices after professional development in these matters. In addition, the research will inform the researcher if the sixth – twelfth grade teachers of the Nonpublic School perceive the implementation of a standards-based reporting system as a change that is significantly different than the norms of the school in which you teach.

If you agree to be in this study, you will be asked to:

- Attend a one-hour informational meeting regarding the study at your school
- Complete an online survey prior to May 28, 2008. This survey will take approximately 20 minutes to complete
- Attend professional development dates already set by the Nonpublic School at the end of the school year. These dates will provide professional development on the use of assessment and grading practices.
- At the conclusion of the professional development sessions, you will have a two-week window to take an online survey on the use of assessment and grading practices. This survey will take approximately 20 minutes to complete.

Your personal information will not be recorded and, therefore, will not be connected to any of your responses. All of your answers will be anonymous. Completing this survey is completely voluntary, and you may quit at any time.

Your participation in this study is voluntary. This means that everyone will respect your decision of whether or not you want to be in the study. No one in the Nonpublic School will treat you differently if you decide not to be in the study. If you decide to join the study now, you can still change your mind later. If you feel stressed during the study you may stop at any time. You may skip any questions that you feel are too personal. If you have any questions, the researcher can be reached at nroorda@aea11.k12.ia.us or by phone at 515.313.1317.

Thank you in advance for your consideration of participating in this research.

Nicole Lynn Roorda

APPENDIX D

Consent Form

You are invited to take part in a research study of the impact of professional development on assessment and grading practices at the secondary level. You were chosen for the study because you are a sixth – twelfth grade teacher in the Nonpublic School. Please read this form and ask any questions you have before agreeing to be part of the study.

This study is being conducted by a researcher named Nicole (Nikki) Lynn Roorda, who is a doctoral student at Walden University.

Background Information:

The purpose of this study is to determine if there is a difference in the use of assessment and grading practices after professional development in these matters. In addition, the research will inform the researcher if the sixth – twelfth grade teachers of the Nonpublic School perceive the implementation of a standards-based reporting system as a change that is significantly different than the norms of the school in which you teach.

Procedures:

If you agree to be in this study, you will be asked to:

- Attend a one-hour informational meeting regarding the study at your school
- Complete an online survey prior to May 28, 2008. This survey will take approximately 20 minutes to complete
- Attend professional development dates already set by the Nonpublic School at the end of the school year. These dates will provide professional development on the use of assessment and grading practices.
- At the conclusion of the professional development sessions, you will have a two-week window to take an online survey on the use of assessment and grading practices. This survey will take approximately 20 minutes to complete.

Voluntary Nature of the Study:

Your participation in this study is voluntary. This means that everyone will respect your decision of whether or not you want to be in the study. No one in the Nonpublic School will treat you differently if you decide not to be in the study. If you decide to join the study now, you can still change your mind later. If you feel stressed during the study you may stop at any time. You may skip any questions that you feel are too personal.

Risks and Benefits of Being in the Study:

Risks and benefits: This research study will ask secondary (6th – 12th grade) teachers to identify their use of academic and nonacademic factors in assessment and grading practices prior to attending professional development on these topics. The researcher will remind all participants

that they have the freedom to withdraw from the study at any time. There are direct benefits to the secondary teachers involved in this study. Benefits include discussion of current assessment and grading practices found within a building; the ability to provide feedback into the assessment and grading practices that will be used in the secondary setting of the nonpublic system; and the opportunity to network with secondary teachers from other buildings during professional development sessions. The secondary teachers' participation in the research may glean a deeper insight into their own assessment and grading practices.

Compensation:

There will be no compensation for the participants of this study.

Confidentiality:

The data collection instrument used in the study will be an online survey that did not ask for demographic information that might reveal the participant's identity. The participants will be assured that their responses, as linked to identity will not be shared with administrators or fellow colleagues. Participants will be informed that the consent forms signed will be kept for seven years and then destroyed.

Contacts and Questions:

The researcher's name is Nicole Roorda. The researcher's faculty advisor is Dr. Howard Carlson. You may ask any questions you have now. Or if you have questions later, you may contact the researcher via the phone at 515.313.1317, or email at nroorda@aea11.k12.ia.us or the advisor via the phone 520.545.2011, or via email at howard.carlson@waldenu.edu If you want to talk privately about your rights as a participant, you can call Dr. Leilani Endicott. She is the Director of the Research Center at Walden University. Her phone number is 1-800-925-3368, extension 1210.

The researcher will give you a copy of this form to keep.

Statement of Consent:

I have read the above information. I have received answers to any questions I have at this time. I am 18 years of age or older, and I consent to participate in the study.

Printed Name of

Participant

Participant's Written or
Electronic* Signature

Researcher's Written or
Electronic* Signature

Nicole Lynn Roorda
nroorda@aea11.k12.ia.us

Electronic signatures are regulated by the Uniform Electronic Transactions Act. Legally, an "electronic signature" can be the person's typed name, their email address, or any other identifying marker. An electronic signature is just as valid as a written signature as long as both parties have agreed to conduct the transaction electronically.

CURRICULUM VITAE

NICOLE LYNN ROORDA

901 SE Sharon Drive
Ankeny, IA 50021
515.963.9511
email: nroorda@aea11.k12.ia.us

EDUCATION:

Doctoral Student 2005 – Present

Focus: Educational Administration for Teaching and Learning
Walden University
Minneapolis, Minnesota

Master of Arts in Education 2001 – 2003 (Honors Graduate)

Focus: Teaching and Learning
Emphasis: Literacy and Curriculum/Assessment
Regis University
Denver, Colorado

Bachelor of Arts 1990 – 1994

Pre-Health Sciences
Elementary Education Certificate
Central College
Pella, Iowa

TRAININGS, CERTIFICATIONS & LEADERSHIP POSITIONS:

Six Plus One Traits of Writing Trainer

Trained Facilitator

Trained in Cognitive Coaching

National Consultant with Heidi Hayes-Jacobs in Curriculum Mapping

Instructional Decision Making Facilitator

Trainer for Collaborative Service Delivery (Co-Teaching/Consultation)

Trained in Fierce Conversations

Trained in Tregoe Analytical Conversations

Learning Focused Relationships

PIE Training

Needs Assessment Training

Iowa Evaluator Approval

PreK-12 Administrator and Supervisor of Special Education
Certification

Balanced Leadership Academy

Project Lead for North Star District Services

North Star Service Delivery Workgroup

Project CRISS

Developing and Implementing the District Career Development Plan

Continuous Quality Improvement Training
 Adjunct Professor for Drake University
 Member of Reading Core School Workgroup
 Member of Collaborative Service Delivery Model
 US West/ISEA Technology Grant Award
 Learning Village Program Leader
 Read In! Coordinator
 Iowa Reading Association Convention Presenter
 Denver Action Research Conference Presenter
 Statewide Literacy Team
 Grade Level Leader
 Building Leadership Team
 Building Literacy Team
 Building Curriculum Team
 Spelling Bee Coordinator
 Nominated for Disney American Teacher Award
 Who's Who in American Teaching
 Douglas County Outstanding Teacher Program
 Cambridge Who's Who Among Executive and Professional Women in
 Education

LICENSURE:

Colorado License Number – 0137272

License – Professional
 Elementary Level
 Elementary Education Specialty

Iowa Folder Number – 327820

License – Professional
 Endorsements - #102 Elementary Teacher, Grades K-6
 #150 Basic Science, Grades K-6
 #101 Athletic Coach, Grades K-12
 #189 PK-12 principal and PK-12 supervisor of
 special education
 Iowa Evaluator Approval License