

Lindenwood University

Digital Commons@Lindenwood University

Dissertations

Theses & Dissertations

Fall 12-2009

Examining Classroom Grades as a Predictor of Student Achievement on a Criterion- Referenced Test: Telling the Truth

Phillip David Kennedy
Lindenwood University

Follow this and additional works at: <https://digitalcommons.lindenwood.edu/dissertations>



Part of the [Educational Assessment, Evaluation, and Research Commons](#)

Recommended Citation

Kennedy, Phillip David, "Examining Classroom Grades as a Predictor of Student Achievement on a Criterion- Referenced Test: Telling the Truth" (2009). *Dissertations*. 591.
<https://digitalcommons.lindenwood.edu/dissertations/591>

This Dissertation is brought to you for free and open access by the Theses & Dissertations at Digital Commons@Lindenwood University. It has been accepted for inclusion in Dissertations by an authorized administrator of Digital Commons@Lindenwood University. For more information, please contact phuffman@lindenwood.edu.

Examining Classroom Grades as a Predictor of Student Achievement on a Criterion-
Referenced Test: Telling the Truth

by

Phillip David Kennedy

December 2009

A Dissertation submitted to the Education Faculty of Lindenwood University in partial
fulfillment of the requirements for the degree of

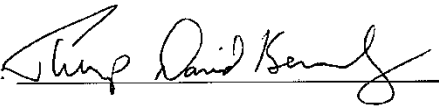
Doctor of Education

School of Education

Declaration of Originality

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

Full Legal Name: Phillip David Kennedy

Signature: 

Date: 12-11-09

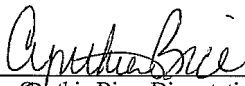
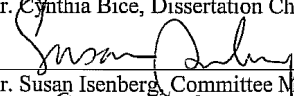
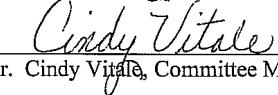
A Dissertation

Examining Classroom Achievement as a Predictor of Student Achievement on a
Criterion-Referenced Test: Telling the Truth

by

Phillip David Kennedy

This dissertation has been approved in partial fulfillment of the requirements for the
degree of Doctor of Education at Lindenwood University by the School of Education.

 _____	<u>12-11-09</u> Date
Dr. Cynthia Bice, Dissertation Chair	
 _____	<u>12-11-09</u> Date
Dr. Susan Isenberg, Committee Member	
 _____	<u>12/11/09</u> Date
Dr. Cindy Vitale, Committee Member	

Acknowledgements

I would like to thank my wife Kelly Kennedy and my sons, Austin and Tyler Kennedy, for their support and understanding while I continued my professional growth throughout the last decade. I would also like to thank Dr. Larry Matthews, Dr. Cynthia Bice, Dr. John Dougherty, and Dr. Susan Isenberg for their help and assistance during the final phase of this journey.

Abstract

American education has made many attempts to reform its structure during the last several decades. Many of these reform efforts have been prompted by global events that implied American education as inferior to other nations' educational systems. The 2001 No Child Left Behind Act required schools to examine their curriculum, instructional practices, and assessments. The problem was a concern regarding inconsistencies between classroom grades and student achievement (Missouri Assessment Program scores). Therefore classroom grades may be misrepresenting student achievement to colleges, military, or corporations recruiting high school graduates.

The purpose of this study was to explore the relationship between student classroom grades and student achievement levels earned through the Missouri Assessment Program (MAP). Classroom grades are symbols and have been used to serve many administrative purposes as well as for feedback on student achievement, instruments for instructional planning, and motivation to achieve more. The elements that comprise the classroom grade were found to include summative and formative assessments, homework, many optional task, as well as grades assigned to non-academic acts such as attendance, behavior, effort and participation.

Eleventh grade Communication Arts data from the study high school were analyzed. A correlation analysis and a chi-square test were conducted using classroom grades and MAP scores. The results from both of these instruments were conflicting. The analysis indicated that there was no significance correlation between the classroom grade and achievement on the MAP test, while the chi-square test indicated there was a

significant relationship. Due to the conflicting results, a further study needs to be conducted using a larger sample size.

Based on the findings of this study, Americans should further examine the classroom grade as a predictor of student success. If the classroom grade continues to be the reporting instrument for student achievement to the public, then the classroom grade should reflect an accurate picture of achievement. A recommendation for future research is to replicate this study in schools that have increased diversity and lower socioeconomic status.

Table of Contents

	Page
Chapter One - Introduction of the Study.....	1
Problem Statement	9
Hypotheses	10
Limitations of the Study	10
Definition of Terms	13
Summary	16
Chapter Two – Review of Literature	18
Research	19
School Leadership	21
Purpose of Grades	24
Grades for administrative purposes	27
Grades used for feedback on student achievement.....	27
Grades for instructional planning.....	28
Grades as motivators	30
Elements of Classroom Grades	32
Summative and formative assessments	33
Homework.....	34
Additional optional task.....	38
Non academic graded acts.....	38
Summary	39
Chapter Three – Research Methodology	41

Participants.....	42
Sampling Procedure	46
Research Setting.....	47
External Validity	48
Research Design	49
Instrumentation	50
Validity and Reliability of the MAP.....	52
The Classroom Grade.....	54
Data Analysis Procedure	56
Summary	58
Chapter Four – Results	60
Analysis	60
2003-2004 Descriptive Statistics	61
2004-2005 Descriptive Statistics.....	65
Discussion	68
Correlation Analysis	71
Deductive Conclusions	75
Comparison of Data Analyses	77
Summary	79
Chapter Five – Discussion	81
Implementations and Recommendations for Effective Schools	86
Change agent	87

Communication	88
Culture	90
Curriculum, instruction, and assessment	92
Relating Findings to O’Connor’s (2002) Model.....	93
Relate grading procedures to the intended learning goals	94
Use criterion-referenced performance standards as reference points to determine goal.....	96
Limited valued attributes included in grades to individual achievement..	97
Sample student performance-do not include all scores in grades.....	98
Grade in pencil-keep records so they can be updated easily.....	100
Crunch numbers carefully if at all	101
Use quality assessments and properly record evidence of achievement..	102
Discuss and involve students in assessments, including grading, throughout the teaching and learning process.....	102
Conclusion	103
Current Status of Grading and Reporting Structure in the Study District.....	104
Recommendations for Future Practice.....	106
Implications of Research Findings.....	108
Recommendations for Future Research.....	109
References.....	110
Vitaé.....	115

List of Tables

	Page
Table 1 – Leadership Responsibilities and their Significance as Measured by Average r scores	23
Table 2 – Reliability Coefficients of the MAP	53
Table 3 – Reliability Information for Educational Assessments Similar to the MAP.....	54
Table 4 – Grading Structures for Teacher A and B Divided by the Percentage Assigned to each Graded Task	55
Table 5 – Achievement Levels as Reported by the Classroom Grade	57
Table 6 – MAP Achievement Levels Divided into Scale Scores.....	58
Table 7 – Comparison Scale for Classroom Grade and MAP Scale Scores	61
Table 8 – Percent of Students Scoring in the MAP Achievement Levels and the Classroom Achievement Levels for the 2003-2004 School year	64
Table 9 – Percent of Students Scoring in the MAP Achievement Levels and the Classroom Achievement Levels for 2004-2005 School Year	68
Table 10 – Mean Classroom Grade and MAP Scale Score 2003-2004 School Year	70
Table 11 – Mean Classroom Grade and MAP Scale Score 2004-2005 School Year	71
Table 12 – Cohen’s Guidelines for Correlation Coefficients	72
Table 13 – Correlation Analysis for 2003-2004 School Year	75
Table 14 – Correlation Analysis for 2004-2005 School Year	76
Table 15 – Leadership Responsibilities and their Statistical Significance as Measured by Average r Scores: An Organized Construct to Create New Meaning.....	84
Table 16 – O’Connor’s Eight Guidelines for Grading in Standards-Based Systems.....	99

List of Figures

	Page
Figure 1 – Gender of student participants for 2003-2004 school year	42
Figure 2 – Student participants with IEP’s for 2003-2004 school year	44
Figure 3 – Gender of student participants for 2004-2005 school year	45
Figure 4 - Ethnic breakdown of student subjects 2003-2005 school year	45
Figure 5 – Student participants with IEP’s for the 2004-2005 school year.....	46
Figure 6 – Ethnic breakdown of Student Participants for 2003-2005 School Years.....	48
Figure 7 – Student subject MAP score by achievement level 2003-2004 school year	62
Figure 8 – Student subject classroom grades for the 2003-2004 school year	63
Figure 9 – Student subject MAP scores by achievement level for the 2004-2005 school year	66
Figure 10 – Student subject classroom grades for the 2004-2005 school year	67
Figure 11 – Scatter plot correlating MAP scale scores to classroom grades for the 2003-2004 school year.....	78
Figure 12 - Scatter plot correlating MAP scale scores to classroom grades for the 2004-2005 school year	79

Chapter One – Introduction of the Study

In the American educational system, students in kindergarten through the twelfth grade receive numerous marks or grades: A, B, C, D, F or Advanced, Proficient, Basic, or Below Basic. These grades signify achievement levels earned during a particular course of study. Beginning in the ninth grade, grades and marks earned by a student are recorded in the students' permanent record. This process continues until the conclusion of the twelfth grade with achievement levels being recorded for each course of study. At the conclusion of the twelfth grade, all of the grades and marks recorded in the students' permanent record are averaged into a numerical statistic that represents the students' over all achievement. This statistic representation known as the Grade Point Average (GPA) is the achievement summation of the students' educational history. This statistic is one of the evaluation instruments for students who enter post-secondary and career education. As one of the evaluation instruments, a question arises regarding the reliability of the GPA.

Letter grades, statistics, and phrases like pass/fail and proficient have been used for more than a century to measure achievement. The use of this traditional grading and reporting system has been a commonly accepted practice. This system summarizes weeks of academic and sometimes non-academic tasks onto a single report, which communicates a student's achievement levels. Despite its common and longstanding use, this type of grading and reporting system has led some to question if it is the most accurate system to use in the grading and reporting of student achievement. A student from The Metropolitan Regional Career and Technical Center in Rhode Island was

quoted as saying, “I’m not a letter in the alphabet. I’m more than that” (Littky, 2002, p. 156).

When the No Child Left Behind (NCLB) Act was passed by Congress in 2001, it represented another era in American educational history. This era, like others before it, ushered in new ideas and the need for change. Educational leaders responded to this need by implementing changes in the areas of curriculum, instruction, and assessment practices. As these changes occurred, educational researchers began to identify a need to change or revise the traditional grading and reporting system. The call to change the traditional system was predicated on the need to more accurately communicate student achievement levels. If educational leaders were already implementing changes in areas that directly impacted student achievement, then it was logical to assume that changes were needed in the grading and reporting system as well. Introducing the idea of another change to the traditional American educational system, however, seems challenging. Change has not always been welcome in the traditional structure. As it is with NCLB, other reform movements in American educational history have attempted to implement improvement and have failed. Despite their failures, many have still left their mark.

Events in American history can have tremendous impact on the American educational system. On Friday, October 4, 1957, Russia launched Sputnik, the first man-made satellite into space. The American people gasped at how a perceived backward nation could complete this task before America. One possible answer was that better schools existed in Russia than in America. According to Bracy (2007), the cover of *Life* magazine for March 24, 1958, read “Crisis in Education.” The story asserted that the average Soviet student was years ahead of even the brightest U.S. high-schoolers” (p.

122). Furthermore, “Not enough math and science were being offered in high schools. Gifted students were being neglected. Other European countries’ systems were superior to our own” (Bracy, p. 127). In response to these concerns, the American government responded by emphasizing the need for more math and science to be taught in American public schools. In addition, Stiggins (1999) asserted, school districts around the country began implementing commercially developed, norm-referenced, district-wide standardized testing programs throughout the 1950s and 1960s in an effort to achieve local accountability (p. 192). This marked the beginning of more reforms to come. In the 1970s, according to Stiggins, the idea of local accountability grew at the state level. By the end of the decade, nearly 40 states were using statewide assessments to report student achievement levels. By 2005, all states were required to have a state reporting assessment as mandated by NCLB.

In the early 1980s, America was embarrassed when the Japanese economy began to prosper while the American economy was sagging. As with the launching of Sputnik, Americans began to point the finger of blame at the educational system. This prompted the National Commission of Excellence in Education to produce a report entitled *A Nation at Risk*, which concluded that America’s public schools were performing at substandard levels. In an effort to raise the bar, Dufour and Eaker (1998) reported that American public schools responded with a variety of strategies that included increasing the number of days that students attended school to increasing the amount of credits students were required to complete in order to graduate. In addition, this era promoted the engagement of administrators in more leadership activities as well as enhanced teacher certification requirements. (p. 3)

Furthermore, educators' attention focused on the types and frequency of student assessments. "A Nation at Risk served as a catalyst for a flurry of school improvement initiatives throughout the United States that came to be known collectively as the Excellence Movement" (Dufour & Eaker, p. 3). Despite the attempt to improve a failing educational system, "The Excellence Movement simply called for an intensification of existing practices. They contained no new ideas" (p. 3).

Throughout this period, not much visible progress was made. The one bright spot could be attributed to the United States Department of Education's implementation of a national assessment program to report student achievement levels. This assessment program is called the National Assessment Program of Educational Progress.

NAEP, or the National Assessment of Educational Progress, is often called the "Nation's Report Card." It is the only measure of student achievement in the United States where you can compare the performance of students in your state with the performance of students across the nation or in other states. (National Center for Educational Statistics, 2009, "What is NAEP?" section)

By the end of the 1980s, American public education was still not producing the type of gains anticipated. The lack of improvements led President George H. W. Bush to initiate a new set of educational goals known as Goals 2000. Goals 2000 was developed with the idea that once fully implemented into America's public schools, the American educational system would finally be where it needed to be in comparison with the rest of the world. Goals 2000 consisted of eight statements that were to be accomplished by the year 2000. These eight goals were as follows:

1. All Children in America will start school ready to learn.

2. The high school graduation rate will increase to at least 90%.
3. American students will leave grades four, eight, and twelve having demonstrated competency in challenging subject matter, including English, mathematics, science, history, and geography, and every school in America will ensure that all students learn to use their minds well, so they may be prepared for responsible citizenship, further learning and productive employment in our modern economy.
4. U.S. students will be first in the world in mathematics and science achievement.
5. Every adult American will be literate and will possess the knowledge and skills necessary to compete in a global economy and exercise the rights and responsibilities of citizenship.
6. Every school in America will be free of drugs and violence and will offer a disciplined environment that is conducive to learning.
7. By the year 2000, the nation's teaching force will have access to programs for the continued development of their professional skills and the opportunity to acquire the knowledge and skills needed to instruct and prepare all American students for the next century.
8. By the year 2000, every school will promote partnership that will increase parental involvement and participation in promoting the social, emotional, and academic growth of children. (North Central Regional Educational Laboratory, n.d., ¶ 4)

Dufour and Eaker (1998) explained that many schools in an effort to reach these goals began implementing a new round of initiatives in an era which came to be known as the Restructuring Movement. The main focus of these initiatives centered on the district level with the implementation of site-based management. Site-based management allowed educational leaders to be creative while helping teachers make improvements in classroom learning. The initiatives and reform ideas of the Restructuring Movement were similar to those that were implemented during the Excellence Movement. Both movements focused on improving educators and schools through programs and requirements that were non-academic in nature. However, because the focus was not on the student, these movements had little impact on student achievement. During the Restructuring Movement, more calls arose for accountability in student achievement. To answer this call, according to Hunt (2008), “Many states began mandating public reporting of student achievement outcomes during this time period” (p. 581).

Furthermore, Hunt (2008) described the next educational movement after the Restructuring Movement as the Standards Movement. During this period of time, attention was shifted from the activities of teachers and administrators to the achievement of students. The Standards Movement emphasized curriculum standards produced by professional organizations. These standards outlined curriculum concepts that students were expected to master in specific subject areas. The standards that were developed were then linked to school improvement activities for those subject areas.

Dufour, Eaker, and Dufour (2004) provided details of the most far-reaching education legislation in modern times. In January of 2002, President George W. Bush signed into law NCLB. This law became the latest attempt for American educational

reform. In the six years since the passage of this law, schools have narrowed their improvement efforts to focus on the areas they are accountable for as defined by NCLB. This law set performance goals to be met according to a time line. If schools did not meet the required performance level, they were still held accountable for meeting the next level the following year. In addition, if schools did not meet the performance goal, they had to operate under specific guidelines outlined in the law. Previous reform efforts never held public educators accountable for results, instead they only hinted at improvement with suggestions and innovations.

Dufour et al. (2004) articulated NCLB's measurement plan. The ultimate goal of NCLB stated that by the year 2014 every student in America would be proficient in Communication Arts and Mathematics. To measure this, schools were evaluated on the basis of a report card known as Adequate Yearly Progress (AYP). The data on the AYP report is broken down into several disaggregated categories to assist districts in identifying those students who did not make the performance goal. The AYP report forced schools to drill down deeper into assessment data to not only identify low performing students but also identify the specific curriculum concepts in which they were low performing. Historically, school leaders would view data to determine where their students were scoring in relation to the state and national average. NCLB brought a new outlook to analyzing assessment data. This new outlook forced administrators to examine the individual classrooms within their district. Before NCLB, schools served as isolated classrooms housed within a single building or campus. At the Missouri School Board Association Conference at Tan-Tar-A in Lake of the Ozarks, October 26, 2002, Dufour

described it in the following way: “They are often one room school houses connected by a common parking lot” (DuFour, 2002, October).

As the face of American education changed, so did individual classrooms. Administrators at a school district in the Midwest began to examine the structure and purpose of the traditional isolated classroom out of a concern that there was a disconnect between the classroom grade and achievement on standardized tests. Curriculum, instructional delivery, and assessment practice were placed under the microscope to determine if they were effective. Curriculum was revised to ensure all teachers were held accountable for teaching grade level and subject area standards. Traditionally, teachers were able to teach the general concepts out of a textbook while highlighting specific areas of interest of that teacher or the students. Teachers generally provided information in a lecture format, and the students received the information. Once the information was received, students were expected to practice and process the information in order to gain a better understanding of the material. This type of instruction generally took the form of answering cogitatively low questions from a textbook or worksheet. At the conclusion of the instructional practice, students were assessed to determine understanding. Teachers often created their own assessments or used one provided by a textbook publisher. Most of these types of assessment never asked students to demonstrate higher cogitative skills by asking them to analyze or synthesize the information.

Due to the perceived low quality of the assessments being used in Communication Arts classes in the study high school, the question remained whether or not the assessments actually measured true student learning. This question opened up more doubts about classroom assessments being used to determine a classroom grade that

eventually was averaged to compute the students' GPA. If classroom grades were representing student learning due to low quality classroom assessments, then the students' GPA may be misrepresenting students' potential MAP test scores.

When NCLB was passed in 2001, school districts were held accountable through the AYP report card that reflected student achievement as measured by MAP scores. In some cases, students were performing well in the Communication Arts classroom as measured by the classroom grade but not performing well on the Communication Arts section of the MAP test. The classroom assessments were comprised of teacher-made tests, unlike the MAP test that was constructed by assessment experts employed by CTB McGraw-Hill. Performance in the classroom was recorded using many factors including averaging of classroom assessments, while the MAP test was being scored by a third party. In other cases, it was the opposite; students were performing well on the MAP but not performing in their Communication Arts classroom as measured by classroom grades. Because of the inconsistency between the two assessment systems, school leaders became concerned that perhaps grades were being inflated or deflated and, therefore, not an accurate summation of classroom student achievement.

Problem Statement

In response to the requirements of the AYP report as outlined by NCLB, schools began examining the instruction and assessment practices carried out by individual classroom teachers. In the examination of these practices, concerns regarding inconsistency between student achievement in the classroom and student achievement on state assessments began to arise. Further, there was a concern that grades were being inflated or deflated by combining non-academic factors with academic factors to

determine classroom grades. The purpose of this study was to determine if there was a relationship between the classroom grade and achievement levels on the MAP test.

Classroom grades from the eleventh grade Communication Arts department at the high school in the study district were compared to eleventh grade Communication Arts achievement levels on the MAP test for the 2003-2004 and 2004-2005 school years.

Hypotheses

H₀: There will be no significant correlation between the Communication Arts classroom grade and Communication Arts criterion-referenced test student achievement levels as measured by MAP scores.

H₁: There will be a significant correlation between the Communication Arts classroom grade and Communication Arts criterion-referenced test student achievement levels as measured by MAP scores.

Limitations of the Study

This study compared eleventh grade Communication Arts students' classroom grades to the same eleventh grade Communication Arts students' MAP test scores. The following limitations were considered.

Sample size. Fifty students from the eleventh grade class of the study high school during the 2003-2004 school year represented only twenty percent of the class. Fifty students from the eleventh grade class of the study high school during the 2004-2005 school year represented twenty-five percent of the class. Because 25% or less of the junior class were represented, the other 75 or 80% of the junior class achievement levels were not introduced in this study.

Sample subject ethnicity. Ninety-eight percent of the 100 sample subjects examined were Caucasian. The other two percent of students at the study high school were from at least five other ethnicities including African American, Asian, Hispanic, Native American, and European.

Geographic location of student samples. The student samples in this study were educated at a suburban high school with an enrollment of 965 students. Because this is a medium size high school according to the Department of Elementary and Secondary Education (DESE), this study would only apply to schools of this size. Data from other size schools may produce different results.

Classroom environment. Student data collected in this study was a result of instruction received in one of two eleventh grade Communication Arts classrooms. Each classroom had its own level of expectations and management established by two different teachers creating inconsistencies between the two. Depending on the learning needs of the student, achievement levels could be imperfect.

Attitude of the subjects. Student subject at the study high school did not perceive the outcome of their MAP achievement level as being important when compared to their grade earned in the classroom. This perception is due in part to the MAP assessment not being connected to graduation or admission to post-secondary institutions. This was confirmed by the central office administration while gathering feedback from focus groups conducted after the MAP testing period.

Socioeconomic status of student samples. Student samples were from the upper and middle class. Only 25% of the entire study high school students qualified under the socioeconomic status category of poverty. According to Payne (1996), students who live

in poverty face more hurdles in their academic journey than students from the middle and upper class.

Consistency of teachers' instructional practice. During the duration of the data collection, students received instruction from one of two teachers. The teachers' implementation of instructional strategies to teach the curriculum concepts was not consistent. Teacher A delivered the curriculum using a lecture/question and answer methodology while Teacher B delivered the curriculum using an experimental learning methodology. The type of instructional delivery, when matched to the students learning style, could impact student achievement.

Consistency of teachers' grading structure. During the time in which the data were collected, students received instruction from one of two teachers. The teachers used different grading structures to assess students' classroom grades. Both teachers also used different variables to determine their classroom grades.

Testing. The classroom grades were a result of assessments created by two teachers. These assessments were different. Both teachers designed assessments that inconsistently covered curriculum concepts. Therefore, these teachers may or may not have covered the curriculum concepts assessed by the MAP test.

Data. Data from the student samples were only examined in the areas of Communication Arts. Teachers of other subject areas were also asked to implement instructional strategies and curriculum concepts to help students increase performance on the MAP test.

Sampling size. This study was limited to sampling only 50 students from a class greater than 200 in each of the study years. Furthermore, this study only analyzed data from the area of Communication Arts.

Definition of Terms

Classroom assessment. Educational tools used to determine student mastery of curriculum content. These tools were administered during and at the conclusion of the study to determine the achievement level earned.

Classroom achievement. This term is used to sometimes describe the classroom grade. The two terms can be used interchangeably when describing the grade a student earned in a course of study.

Classroom grade. The classroom grade is a final letter grade assigned to represent the students overall achievement for the course of study. The grade is represented in the form of an A, B, C, D, or F. Each letter grade signifies the degree to which each student mastered the courses curriculum concepts. The classroom grade is comprised of both academic and non-academic factors.

Criterion referenced assessment. This test is one that serves as a reference point by comparing performance standards questions answered correctly to students' previous achievement on the same performance standards questions. Criterion referenced tests compare data from performances of the same student (Popham, 2001).

Data analysis. Analyzing classroom assessment outcomes to determine student understanding of curriculum concepts during a unit of study. This analysis is used to design instruction to meet the needs of each student.

Department of Elementary and Secondary Education (DESE). Governmental department responsible for monitoring 524 public school districts in the state of Missouri. DESE conducts reviews of Missouri school districts through a review process known as Missouri School Improvement Process. This process is conducted once every five years to help school districts identify strengths and areas of improvement (Missouri Department of Elementary and Secondary Education [DESE], 2004).

Formative assessment. An assessment given during instruction to collect information on student understanding. This information serves as a report to the teacher to identify which students are, or are not, understanding the curriculum concepts being taught. In addition, Stiggins, Arter, Chappuis, and Chappuis (2006), concluded that formative assessments were used to inform students about themselves. Formative assessments can be formal or informal. Formative assessments need to be conducted on a frequent basis.

Grade Point Average (GPA). The GPA is the combination of multiple courses classroom grades. Each letter grade is assigned a numeric representation. These numbers are then averaged to calculate the students' overall average achievement for the four years spent in high school.

Missouri Assessment Program (MAP). An education reform mandated by the Outstanding Schools Act of 1993. This reform identified the knowledge, skills and competencies that Missouri students should demonstrate mastery of by the time they complete high school. The MAP test is a criterion-referenced assessment given to Missouri students during the spring testing window. This assessment is given in grades 3-

8 and 11 in the subject area of Communication Arts and grades 3-8 and 10 in Mathematics (DESE, 2004).

No Child Left Behind (NCLB). An education reform enacted into law in the year 2002. This reform was designed to measure progress toward ensuring that all students in the United States performed at pre-determined levels of achievement by the year 2014. NCLB measures student achievement in the areas of Communication Arts and Mathematics (DESE, 2008).

Standard-based assessment. This test is one that serves as a reference point by comparing performance standards answered correct to other students' performance on the same test.

Standardized test. Norm and criterion referenced test designed to measure student achievement. This type of test addresses different curriculum areas using multiple choice and essay questions. Student scores are often used for placement in academic programs or as predictors of academic success (Popham, 2001).

Student achievement. Academic levels earned by students by meeting requirements outlined by the classroom grade and MAP test. Students will score at one of five achievement levels in the classroom (A, B, C, D, and F) and one of five achievement levels (Advanced, Proficient, Nearing Proficient, Basic, and Below Basic) on the MAP test.

Summative assessment. An assessment given at the end of an instructional period to determine student mastery over curriculum content. This assessment encompasses questions that cover all the curriculum concepts at a point in time. Summative

assessments are used to report student achievement to third parties (Chappuis, Stiggins, Arter, & Chappuis, 2004).

Teacher collaboration. Teachers working together to achieve a common goal. Dufour and Eaker (1998) reported that “creating small, supportive groups in which teachers are encouraged to discuss their questions, concerns, and ideas” about student learning enhances the rate of success (p. 6). This collaboration takes place during specified times and generally involves the sharing of a teacher’s practice in the classroom.

Summary

During the past several decades there has been a perception that American education has never measured up to that of other civilized nations around the world. This perception is based on events in history such as the Russians launching the first man made satellite into space and the Japanese economy outperforming the American economy. Because other nations were able to outperform America, the finger of blame pointed to its educational system. The response by educators was summarized by reform initiatives known as the Excellence Movement and the Restructuring Movement. Both attempts at restructuring the American educational system were unsuccessful due to the revamping of old initiatives and failing to provide any new ideas. In 2001, NCLB was signed into law. NCLB brought a fresh concept into American education: accountability. Because of NCLB, districts were accountable for student performance goals in the areas of Communication Arts and Mathematics. As a result, educational officials reexamined and revised curriculum and assessments. In addition, educators were mandated to evaluate the instructional practices that were taking place in the classroom.

Because of NCLB, educators started to focus on teaching practices. Teachers began to collaborate with one another, unlike they had done in the past. Through all of the curriculum revision, collaboration, and data analysis, a concern began to arise, centered around the relationship between the student classroom grade and student performance on standardized tests. It appeared that while schools were producing a large number of graduates each year, they were not meeting the MAP test performance targets mandated by NCLB. Chapter two will review the study's associated literature.

Chapter Two - Review of Literature

Since the introduction of NCLB, the American government has sent the message to the public that American education is a priority. In fact, it seems that recent United States presidents have put education at the top of their list. Danielson (2002) believed the individual reasons may vary, but the main motivation most likely stems from the issue that, with few exceptions, student achievement has not significantly improved. Waters, Marzano and McNulty (2005) reported that school leadership had a significant impact on student achievement. For this reason, the twenty-one school leadership responsibilities that impact student achievement were examined. In addition, O'Connor's (2002), eight guidelines for grading were reviewed to better understand the impact they had on student achievement.

Several factors related to the classroom grade were examined to determine if these factors contribute to overall reporting of classroom achievement. Classroom achievement has usually been reported in the form of a letter grade of A, B, C, D, and F. In American education, grades have always been a gauge for measuring student achievement. Because this is the measurement, the many factors that comprise the classroom grade were examined to determine their role when assigned to the classroom achievement levels. The review of the literature examined the historical perspective of the classroom grade as well as the factors that contribute to the classroom grade. In addition, it examined the elements that typically make up the classroom grade. These elements included summative and formative assessments, homework, additional optional tasks, and non-graded tasks.

Research

According to Olson (1995), for the past hundred years, the grades assigned by teachers were used to represent the student's classroom achievement. In addition to providing achievement data to the student, it also provides achievement data to third parties as well. Classroom grades, as achievement indicators, have been the primary tool for communicating to parents and post secondary institutions on how the student is mastering curriculum concepts.

Despite this being the primary measurement tool for decades, controversy has surrounded this method for as long it has been used. The researchers Cross and Frary (1999) illustrated, "School marks and grading have been the source of continuous controversy since the turn of the century" (p. 55). Cross and Frary documented the concern of Finkelstein in 1913 in the use of marks as an uncalibrated instrument. Furthermore, Middleton (1933), Guskey (1996), and Marzano (2000) stated their concerns about the use of grades as a measurement instrument. Despite the various degrees of concerns, most researchers agreed on one or more of the following three factors: (a) teachers consider many factors other than academic achievement when they assign grades, (b) teachers weigh assessments differently, and (c) teachers misinterpret single scores on classroom assessments by mixing different types of knowledge and skills into single scores on assessments.

Marzano (2000) conducted an analysis of the factors that impacted grading. He concluded that a common element based on the latitude and freedom that the American education system allows in its measurement system created misguided information. Grading had become based on the subjectivity and inconsistency of educators' decisions

in determining the grade. Due to the inconsistencies of grades, Marzano (2000) stated they have become so imprecise that they are almost meaningless. Guskey (1996) described

The one reason educators are inconsistent in grading practices can be linked to lack of proper training and practice in state mandated certification programs. This lack of training and practice means educators generally rely on grading practices that are not the result of careful thought or sound evidence; rather, they are used because teachers experienced these practices as students and, having little training or experience with other options, continue to use them. (p. 18)

Not only are there inconsistencies among educators in determining grades within their own classroom, inconsistencies are also prevalent among educational institutions. Because different educational institutions have different standards, some might define their achievement levels differently. In a traditional grading structure, where achievement levels of A, B, C, D, or F are used, the manner in which one institution defined an achievement level was different from the same level defined by another institution. For example, if a student receives a grade of an A, does that mean they mastered 97% of the content or 95% of the content? Can students pass a course having mastered 60% (D) of the content, or do they need to have mastered 62% (D) or perhaps 70% (C)? Marzano (2000) pointed out that the discrepancy among educational institutions initiated the question regarding the precise definition of grades, leading to the creation of the concept of grade inflation.

School Leadership

When it comes to inconsistencies in American grading structures, the element of school leadership can correct the structure. Leadership can be the deciding factor in determining success or failure. In addition, leadership has been the determining factor in the outcome of many historical events. Just as is the case with education, leadership is the key element in determining success or failure for schools. The mission of every American school is to provide an education to its students so they can be successful once they graduate. Therefore, it is the responsibility of school leaders to provide the structure necessary for their schools to be effective and reflect accuracy in the many components of education. The component of grading and grade reporting is one where this is necessary.

Educational leaders must understand the need for revising the grading and reporting structure. In addition to understanding the need for structural revision, educational leaders need to know how to implement this revision in their schools. To help educational leaders identify the steps necessary to revise the grading and reporting system in their school or district, the model suggested by researchers Marzano, Waters, and McNulty (2005) of the Mid-Continent Regional Educational Laboratory (McREL) was applied to this process. These researchers identified 21 leadership responsibilities related to student achievement. These responsibilities identified the impact of school leadership on student achievement. The analysis of this study identified a new leadership framework that was more comprehensive than previous frameworks. The reason for this was that it was grounded in more concrete evidence based on thirty years of research. In addition, this leadership framework was predicated on the notion that “effective leadership means

more than simply knowing what to do, it's knowing when, how and why to do it" (Marzano et al., 2005, pp. 60-61). Marzano et al. (2005) believed "effective leaders must understand how to balance pushing for change while at the same time protecting aspects of culture, values, and norms worth preserving" (p. 45). These 21 responsibilities represented a balance of knowledge and skills that leaders must exhibit to positively impact student achievement. The impact is measured by the Average r . The Average r is the range of correlation coefficient. The range of the Average r extends from -1 to 0 and 0 to 1. It represents to what degree a relationship exists between the leadership responsibility and student achievement. The closer the Average r is to 1, the stronger the relationship. Table 1 represents the responsibilities and their related Average r .

Table 1
Leadership Responsibilities and Their Significance, as Measured by Average r Scores

<u>Responsibility</u>	<u>Average r</u>
1. Affirmation	.19
2. Change Agent	.25
3. Contingent Rewards	.24
4. Communication	.23
5. Culture	.25
6. Discipline	.27
7. Flexibility	.28
8. Focus	.24
9. Ideals Beliefs	.22
10. Input	.25
11. Intellectual Stimulation	.24
12. Involvement in Curriculum, Instruction, and Assessment	.20
13. Knowledge of Curriculum, Instruction, and Assessment	.25
14. Monitoring/Evaluating	.27
15. Optimizer	.20
16. Order	.25
17. Outreach	.27
18. Relationships	.18

19. Resources	.25
20. Situational Awareness	.33
21. Visibility	.20

Note. From *School Leadership that Works: From Research to Results* (pp. 42-43), by R. J. Marzano, T. Waters, and B. A. McNulty, 2005, Alexandria, VA: Association for Supervision and Curriculum Development.

After the 21 leadership responsibilities were identified, they were correlated to the impact they had on student achievement. The correlation Average r score of the twenty-one responsibilities was found to be .25 (Marzano et al., 2005). A correlation average of 1 is a perfect correlation, though some researchers do not recognize .25 as being significant, as did the McREL researchers. Table 1 identifies each individual leadership responsibility as well as the Average r associated with it. The closer to 1 the Average r number, the greater the significance.

Leadership is the key to whether a school will be either effective or ineffective. Through the research of Marzano et al. (2005), the implementation of the 21 leadership responsibilities in a school will have an impact on student achievement. Therefore, these qualities must be present if achievement as defined by the classroom grade and achievement on the MAP test are accurately reported.

Purpose of Grades

Classroom grades are a combination of many factors that contribute to a student's learning process. Because of this, many researchers have come to define the actual purpose of grades. According to Guskey (1996), there are five purposes for grades: (a) to

communicate the achievement status of students to parents, students, and others; (b) to provide information that students can use for self-evaluation; (c) to select, identify, or group students for certain educational paths or programs; (d) to provide incentives to learn; and (e) to evaluate the effectiveness of instructional programs. Brookhart (2004) believed the primary purpose for grading should be to communicate with students and parents about students' achievement of learning goals. The secondary purpose for grading includes providing teachers with information for instructional planning and providing teachers, administrators, parents, and students with information for placement of students. According to Birk (2000), teachers tend to give grades for many different reasons: (a) to measure content mastery, (b) to chart progress, (c) to motivate students, and (d) to provide information to a variety of audiences from students to parents to college admission boards. Finally, Airasian (1994) explained that educators use grades primarily for the following five reasons: (a) for administrative purposes, (b) to give students feedback about their progress and achievement, (c) to provide guidance to students about future course work, (d) to provide guidance to teachers for instructional planning, and (e) to motivate students. Whichever definition is used, according to Birk (2000), it is clear that most researchers agree that the purpose of grades is to provide information and feedback about student achievement to students, parents, and others. In addition to the above mentioned purpose of grades, some educators also suggest that classroom grades are a predictor of future academic success.

Because grades are the instrument used to provide information and feedback to third parties, they must represent an accurate picture. Assessment specialists including Stiggins, McTighe, and Guskey indicate that seven underlining perspectives on grading

have been developed (as cited in O'Connor, 2002). The seven perspectives are as follows:

1. Grading is not essential for learning.
2. Grading is complicated.
3. Grading is subjective and emotional.
4. Grading is inescapable.
5. Grading has limited research base.
6. Grading has no single best practice.
7. Grading that is faulty damages students and teachers. (p. 17)

Because of these perspectives, O'Connor presented eight guidelines to support learning and to encourage student success. The eight guidelines are as follows:

1. Relate grading procedures to the intended learning goals.
2. Use criterion-reference performance standards as reference points to determine grades.
3. Limit the valued attributes included in grades to individual achievement.
4. Sample student performance - don't include all scores in grades.
5. "Grade in pencil" - keep records so they can be updated easily.
6. "Crunch" numbers carefully - if at all.
7. Use quality assessment(s) and properly record evidence of achievement.
8. Discuss and involve student in assessment, including grading, throughout the teaching/learning process. (pp. 243-244)

These eight guidelines can be examined in relation to the purpose of grades.

Grades for administrative purposes. Durm (1993) discussed the history of using grades for administrative decision-making when determining placement, advancement, and retention of students. Grades could be used for placement into courses when one student transfers from one school to another. This researcher, as the Assistant Superintendent of Curriculum and Instruction for the study district, notes that at the secondary level, grades in particular courses are also used to determine if a student is ready to enter the next sequential class. Grades continue to be used to determine advancement beyond secondary education as post-secondary institutions review student grades as well as achievement level (grade) earned on the college entrance exam(s). These two measurable indicators are the major factors in determining if a student is accepted into a post-secondary institution. In relation to O'Connor's (2002) guideline number two, the meaning of a grade should come directly from the standards set forth by the course in which the grade was earned. Because grades are used for the above mentioned reasons, they should also communicate the course standards that the student mastered.

Grades used for feedback on student achievement. Grades are also used to give students feedback about their progress and achievement. Durm (1993) reported this use of grades was traced back to 1780 when Yale University began using a system based on a four-point scale. Soon afterwards, educational institutions such as William and Mary University, Harvard University, and many others adopted their own versions of the grading scale. Again, the debate among researchers and educators raged regarding whether or not a single letter grade representing multiple facets of learning constituted accurate feedback.

Brookhart (2004) posed the question of whether a letter grade represents a student's growth in a content area or whether learning had taken place. Aspects such as educational background and training contributed to educators' interpretation of a letter grade to mean growth or demonstration of learning. To some, it was the student's point total for the classroom assignments and assessments. For others, it was that point totals plus other aspects of contribution such as effort or attitude. O'Connor's (2002) guideline number three maintains that "the classroom grade should only represent academic achievement. Other valued attributes such as behavior and attendance should not be communicated as part of the grade" (p. 89). Furthermore, guideline four advocates not including all scores in the final grade. Feedback should be given by using a variety of assessments and formative performance products. Because students are processing their learning through practicing methods, they should not be graded on everything. "The main difficulty driving grading issues both historically and currently is that grades are pressed to serve a variety of conflicting purposes" (O'Connor, p. 89).

Grades for instructional planning. In the age of accountability as established by NCLB, educators began to take a closer look at grades and analyze their true meaning. Prior to this accountability in education, little time was spent on examining the representation of grades. Traditionally, educators gave students a letter grade which, they felt, was earned by completing a series of classroom tasks. Once a letter grade was recorded for completion of these tasks, teachers moved on to the next task. With teachers simply recording grades and moving forward, there was no explanation for what the grade of an A, B, C, D, or F actually meant with regards to a student's progress. Popham (2001), reflecting on his experience as a classroom educator, said

When I was a high school teacher, using tests to make instructional decisions never even entered my consciousness. Along with my fellow teachers, I used tests for one purpose only: to assign grades. Even today, the majority of teachers employ tests either for grade-dispensation (if you scored higher than 95% on the midterm exam, you will get a grade of A). (p. 32)

Educators looked for consistency in grade representation by analyzing grade data. This analysis caused them to determine the ultimate number of curriculum concepts that must be mastered to justify a certain recorded letter grade. Prior to this grade investigation, data analysis in reference to the individual classroom grade had not taken place out of fear. Schmoker (1999) recognized educators were fearful of data because of its capacity to reveal strengths and weaknesses, failure and success. Being fearful of data and thereby ignoring it promoted inaction and inefficiency. There is very limited evidence that suggests data such as grades were for instructional planning purposes in the past. The accountability phase in education forced educators to use multiple data sets for instructional planning. For instance, data began being used to pinpoint specific subject areas where students are not mastering curriculum concepts, forcing educators to go back to reteach those areas until mastery can be recorded. Schmoker (1999) further noted

Data can help us confront what we may wish to avoid and what is difficult to perceive, trace, or gauge; data can substantiate theories, inform decisions, impel action, marshal support, thwart misperceptions and unwarranted optimism, maintain focus and goal-orientation, and capture and sustain collective energy and momentum. (pp. 48-49)

Grades as motivators. As described by Guskey and Bailey (2001), grades are also consistently embraced as motivators for student achievement. “If students are motivated to learn the content in a given subject, their achievement in that subject will most likely be high. If students are not motivated to learn the content, their achievement will be limited” (Guskey & Bailey, p. 35). The reality of grades being a motivator has often been debated. According to Blount (1997), 82% of teachers said they used grades to motivate students. One teacher in Blount’s study said, “It isn’t because the grades motivate the students to work hard. I believe that if we were on a pass/fail system most students would do the bare minimum to pass” (p. 330). Marzano (2003) reported motivational theory is complex, involving multiple dynamics such as (a) drive theory, (b) attribution theory, (c) self-worth theory, (d) emotions, and (e) self-system.

Drive theory is characterized by two driving forces: “Striving for success and the fear of failure” (Marzano, 2003, p. 145). Students who are motivated by the drive theory are motivated by emotional rewards or not motivated to do new tasks because failure incurs a negative effect. For example, students may work to earn the achievement level of an A because they strive for success; their expectation level is set very high. Other students will continue to work on curriculum concepts and go the extra mile because they want to ensure that they have done everything and more to earn a high achievement level.

Attribution theory is defined as “success or failure is based on ability, effort, luck, or task difficulty” (Marzano, 2003, p. 146). Students who are motivated in this manner generally use past experiences as their reason for being motivated or not motivated. For example, if a student experienced academic success due to devoting a great deal of time to studying, they might continue to practice this method in the future. Conversely, if a

student spent minimal time studying and earned a high achievement level, they may not find it necessary to study for future assignments.

Self-worth theory is based on “self-acceptance as one of the highest priorities in one’s immediate or peer culture” (Marzano, 2003, p. 146). In the boundaries of the self-worth theory, students may confuse ability with worth. Students who are motivated by self-worth means often will ask for continuous feedback or additional opportunities to please themselves and others. The students’ work ethic is driven by being accepted by those around them.

In the self-system theory, students’ needs and aspirations are organized into a hierarchical structure similar to that outlined by the work of Maslow. Maslow’s (1943) hierarchy of needs is based on five levels, the lowest fulfillment of one’s basic needs to the highest level of self-actualization. The needs are structured in a manner in which a person can only move through one level at a time. In addition, a person cannot proceed to the next level until the level they are currently at is complete.

Whichever motivation theory a student works within, grades can be applied to any of these theories. Grades are a factor that may cause students to strive to do their best or may be the reason why students are discouraged from demonstrating what they have learned. Guskey and Bailey (2001) stated

Grades are sometimes viewed as tools of reward and punishment by teachers.

Teachers consider grades or reporting forms their ‘weapon of last resort’. In their view, students who do not comply with their requests suffer the consequences of the greatest punishment a teacher can bestow: a failing grade. Such practices have

no educational value and, in the long run, adversely affect students, teachers, and the relationship they share. (p. 35)

Furthermore, Guskey and Bailey reported that “no studies support the use of low grades or marks as punishments. Instead of prompting greater effort, low grades more often cause students to withdraw from learning” (p. 35).

O’Connor’s (2002) guideline five suggests that teachers grade in pencil. This implies that students are given several opportunities to demonstrate learning. The more opportunities presented, the greater the chance a student has to increase their grade. If students know they have an opportunity to raise a low grade, they may be motivated to continue trying. If the low grade is earned and no other opportunity is presented to correct it, then the student may not be motivated to learn any further for fear of continued failure.

Similarly, Littky (2002) believed the classroom grade has not always been a true representation of a relationship between motivation and learning. If this is true, then students who work hard would receive the highest grade. But because hard work does not necessarily translate into understanding, the grades they earn may not match the effort they put in. For example, if a student was motivated to master the curriculum concepts by completing all of the assignments and comes away earning the letter grade of C, the final grade of a C may not accurately reflect the effort put forth by the student. In addition, because the student’s final grade does not reflect the effort invested, it may deter the student from putting the same amount of effort toward other curriculum concepts. In the world outside of education, employees are not graded on the job. They receive feedback about their performance and are told what they need to do to improve.

Elements of Classroom Grades

A child attends school to learn. As a report of their learning, children earn a letter grade. This report may not have necessarily been a reflection of learning or the effort put forth in the learning process. In fact, according to Bailey and McTighe (1996), “Grades often reflect a combination of achievement, progress, and other factors. The tendency to collapse several independent elements into a single grade may blur its meaning” (p. 119). In the past, students may have felt that the effort and work they put into a task was not reflected in the grade they received. Or, perhaps the letter grade was inflated by the classroom teacher based on other variables, which may or may not have been related to the students’ demonstration of learning.

There are many variables that teachers often include in their grading structure that can either hurt or enhance a student’s letter grade. Some of the more commonly used variables include (a) assessments (formative and summative), (b) homework, (c) additional optional tasks (extra credit), (d) attendance, (e) behavior, and (f) effort and participation. The first three variables, assessments, homework and optional tasks, are typically associated with academic achievement. The last three, attendance, behavior and effort, are non-academic and more closely tied to classroom management. Marzano (2000) believed because classroom management is a vital part of education, many educators tend to include those factors into their overall classroom grading policies.

Summative and formative assessments. Classroom assessments have taken many forms in American education. Two of the most common forms have been the summative and formative assessments. According to O’Conner (2002),

A summative assessment is one that assesses how much a student has learned over a period of time. It is an assessment or evaluation designed to provide information to be used in making judgments about a student's achievement at the end of a sequence of instructions, e.g. final drafts/attempts, test, exams, assignments, projects, performances. (pp. 109-110)

Furthermore, Stiggins et al. (2006) defined summative assessments as assessments that demonstrate of learning. The learning is a reflection of multiple curriculum concepts reported over a longer period of time. Due to the nature of the definition, these assessments resemble a test that is given at the conclusion of a unit of study. In the traditional classroom, educators have relied heavily on the use of summative assessments. At the conclusion of the summative assessment, students were assigned a grade, the next topic was introduced, and the process was repeated.

O'Connor (2002) described formative assessments as
Assessments that are given more frequently throughout the unit of instruction. They are assessments designed to provide direction for improvement and/or adjustment to a program for individual students or for a whole class (e.g., observation, quizzes, homework, instruction questions, initial drafts/attempts).

This type of assessment informs both teachers and students about student understanding at a point when timely adjustments can be made. (p. 113)

Garrison and Ehringhaus (n.d.) believed that "formative assessment helps the teacher determine the next steps during the learning process" (p. 2). For many educators, formative assessments merely represent practice for the students on their journey before engaging in the summative assessment. Because formative assessments are being used for

practice, teachers are not assigning grades or marks to them. Chappuis et al. (2004) described formative assessments as assessments for learning, designed so that educators can determine how they can help students learn more during the unit of study.

Checking for understanding and re-teaching for mastery were never given much attention. But with the resurgence of formative assessments and data analysis due to NCLB's accountability requirement, they have come to the forefront. Educators are now using classroom assessments to make inferences about student learning. Typically, a child is tested to determine what the child knows or can do, but tests are also being administered to assess a student's attitude or interest. Teachers need this kind of information to make sensible instructional decisions about their students and curriculum. Popham (2001) stated the chief mission of any classroom test should be to capture the kind of information teachers need to make better instructional decisions. Classroom assessments are more than an assigned letter grade. They are tools designed to give the teacher feedback on whether or not students understand and have mastered the content. In addition, formative assessment are now being used to help inform the teacher about instructional decisions that need to be made in order to meet the needs of all students.

Homework. According to Fraser, Walberg, Welch, and Hattie (1987), "Students typically spend anywhere from six to eight hours a day at school. It occupies about 13 percent of the waking hours of the first 18 years of life" (p. 234). During this time, they are engaged in multiple learning activities in various content areas. A student's day, five days a week, is spent being exposed to new learning opportunities. In addition to all of the learning that takes place during the school day, educators feel the need to assign additional work to be completed at home. The assigning of homework has been a practice

that has been commonplace for many years. It is taken for granted that this practice is an expectation of teachers, administrators, and parents. Despite this expectation, many have debated the impact that assigning of homework has on learning.

In order to examine the impact that homework has on learning, one must first examine what defines homework. According to Marzano, Pickering, and Pollock (2001), “Homework and practice are instructional techniques that are well known to teachers. Both provide students with opportunities to deepen students’ understanding and skills relative to content that has been initially presented to them in class” (p. 60). Marzano et al. (2001) described homework as an opportunity to extend learning opportunities beyond the confines of the school day. When a student is assigned homework, it should be based on information that they have already been taught. The homework is for students to practice with that learned information or to elaborate on the concepts introduced. In theory, the homework strategy is one that could prove to have a positive impact on student learning. Kohn (2006) argued that homework was not used in an appropriate manner and therefore the long-standing educational practice of assigning homework has little to no impact on student learning. Cooper (1998, as cited in Kohn, 2006), confirmed this theory in a study conducted with both younger and older students (grades 2-12). This study revealed that there was no significant relationship between grades achieved and the amount of homework assigned. When the concept of homework completed was compared to the effect it had on grades, Cooper found that there was a positive relationship with the older students, grades 9-12. Cooper’s study also concluded that there was a negative relationship between homework completed and the effects it had on grades with the younger students.

It seems strange that such a common practice of assigning homework would continue to exist if there is evidence that it does not positively impact student learning. There are several other factors that may support the theory that homework may not be an effective teaching practice. When homework is taken outside the classroom, it may be completed with the assistance of a second party. If the student did not fully understand the initial learning that took place at school, they may not be able to complete the homework at home, causing the student to ask for assistance or not do the assignment at all. If assistance by a second party is provided, then that person may not fully understand the intended purpose of the assignment or may not possess the knowledge needed to accurately assist the student. The student may copy the assignment from someone who had already completed it. If this occurs, then the student did not receive the opportunity to be engaged in the extended learning. The student is only motivated to complete the assignment for the impact it would have on his/her overall classroom grade. The student has little interest in demonstrating mastery of the curriculum concepts. If this process is repeated by the student, the final grade will not be a true representation of the students' overall learning. The grade then becomes a representation of being able to turn in assignments and receive points not an indication of learning.

The United States Department of Education (2003) has stated, Homework should only serve one of four purposes: (a) Practice- to reinforce learning and help students master specific skills; (b) Preparation- introduces material presented in future lessons. The assignments aim to help students learn new material when it is covered in class; (c) Extension- asks students to apply skills they already have in new situations; and (d) Integration- requires students to

apply many different skills to a large task, such as book reports, projects, creative writing. (p. 2)

If homework is given for purposes other than the four stated above, it could have a negative impact on student learning.

Additional optional tasks. Many times at the conclusion of a grading period, a student who may have aspired to earn a higher letter grade than what will be recorded is motivated to find alternative ways to raise his/her grade. When this occurs, students typically approach the teacher about ways in which their grade can be raised. Teachers who do not wish to see their students fall below expectations oblige the students request by offering additional tasks to complete before the end of the class term. These additional tasks include making up work that was not previously turned in on time. Sometimes, teachers offer students bonus points or extra credit for turning in an additional writing assignment or a project that covers a concept taught previously during the grading period. This option gives some students an opportunity to earn points that the rest of the students were not given. The completion of additional and optional tasks leads the student to focus only on accumulating enough points to earn the desired grade. The points accumulated in no way reflect whether or not learning or understanding of content took place in the classroom.

Non- academic graded acts. The intended purpose of a classroom grade is to be a representation of what a student learned. As described by Cross and Frary (1999), there are many other non-academic factors included in a student's final grade. In research conducted by Cross and Frary, 39% of teachers admitted including non-academic factors in grades. These non-academic factors such as attendance, behavior, effort and

participation were often factored into classroom grades because many thought they contributed to whether or not a student had been provided with the opportunity to learn. As reported by Brookhart (2004), teachers perceive the three “non achievement” factors of effort, behavior, and attendance as important to classroom control and, consequently, often include them in their grading policies.

Parents and perspective employers would agree that developing non-academic factors is necessary, but others would argue fairness to assess these in conjunction with academic criteria cannot be justified. In schools throughout the United States, the concepts of citizenship and character are implemented into a student’s daily curriculum because so many school districts have recognized the importance for students to possess these skills to be productive members of society. It is important for schools to provide feedback to students on non-academic concepts if they include them in their daily curriculum. Marzano (2000) stated feedback from non-academic concepts should be kept separated from the feedback given by the academic grade.

Summary

The purpose of classroom grades is to provide feedback to students, parents, and others about the academic achievement of a student. Classroom grades have been used for administrative purposes as well as guidance in instructional planning. In addition, classroom grades also serve as a motivator, both positive and negative for students. Despite the purpose of the classroom grade, it has been a representation of student achievement for over one hundred years. During this time, many researchers have come to doubt the validity of the classroom grade as a true measurement of student achievement. Grades have been accused of being an imperfect representation of student

achievement because of the various factors that are often included in the end result calculations. The factors are often a combination of academic and non-academic tasks. In many cases, the classroom grade did not clearly tell the entire story about a student's achievement over a period of time due to the latitude and freedom practiced by educators. Chapter three examines the methodology used in this study to examine the relationship between classroom grades and achievement on state mandated tests.

Chapter Three – Research Methodology

In American education, report card day is either eagerly anticipated or dreaded. For students expecting an exemplary report card, it is viewed as a sense of accomplishment. For those students who expect a report card with unsatisfactory marks, the day is one to avoid. To parents, the report card is the official document that communicates their child's level of achievement. The grades are the representation of the amount of intellectual growth that took place during a given time period. Parents, by and large, have accepted the classroom grade as an assessment of a particular mastery level. Rarely have parents taken the time to truly examine the classroom grade and how it is determined. They continue to accept the grade instead of examining what the grade measures. Because the classroom grade often measures elements other than academics, it may not be a true measure of academic achievement. This has created confusion when a child's grade does not match the performance as reported by an assessment measuring only academic achievement. This relationship has become a topic worthy of examination to determine if academic achievement on a standardized test is a similar representation of a classroom letter grade.

In 1994, DESE contracted with CTB McGraw Hill to develop a grade span assessment. This grade span assessment became the MAP. The MAP test is a standardized assessment designed to measure academic achievement of Missouri students in relation to the Show-Me-Standards.

Participants

This study compared the classroom grade of 50 junior English III students at the study high school to their achievement levels earned on the eleventh grade

Communication Arts MAP assessment during the 2003-2004 school year. In addition, this study examined the same achievement levels for an additional 50 students during the 2004-2005 school year. A total of 100 student achievement levels were analyzed. This study attempted to determine if there was a correlation between the achievement levels reported by the classroom grade and the achievement level reported on the MAP.

As illustrated in Figure 1, during the 2003-2004 school year, 50 student subjects were randomly selected from the junior class at the study high school to examine achievement levels. Of these 50 subjects, 24 were male and 26 were female. Although the study high school did have a small percentage of students from other ethnic backgrounds enrolled in the junior class during the 2003-2004 school year, all of the 50 student subjects were Caucasians.

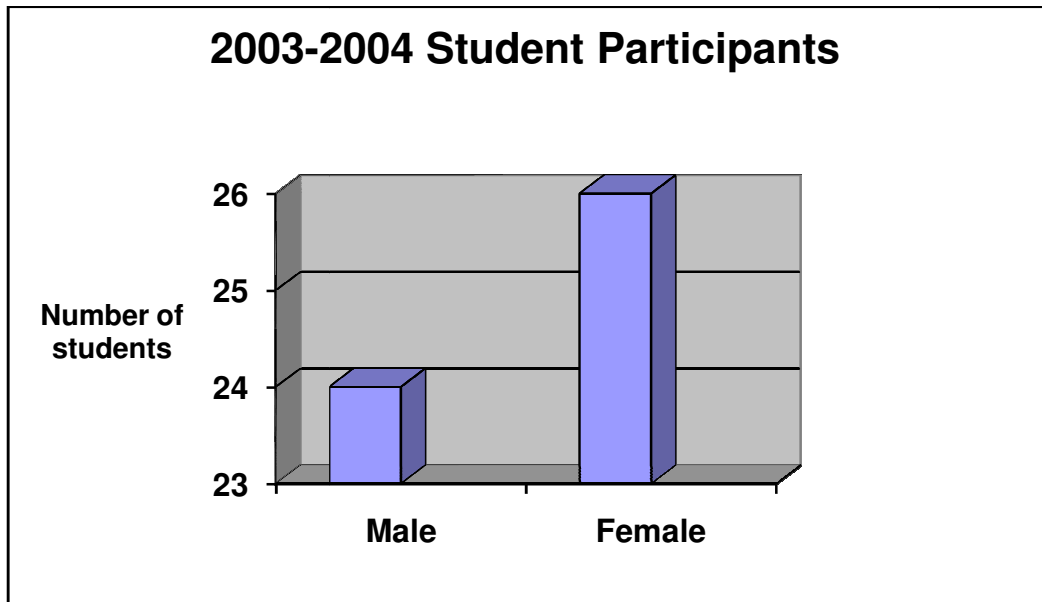


Figure 1. Gender of student participants for 2003-2004 school year.

This study also examined the number of student subjects who were receiving additional educational services. These include additional services in the form of another classroom teacher providing assistance or the student being pulled out of the classroom for one-on-one instruction. Three of the 50 randomly selected student subjects received additional educational services as documented in their Individual Educational Plans (IEPs) as illustrated in Figure 2.

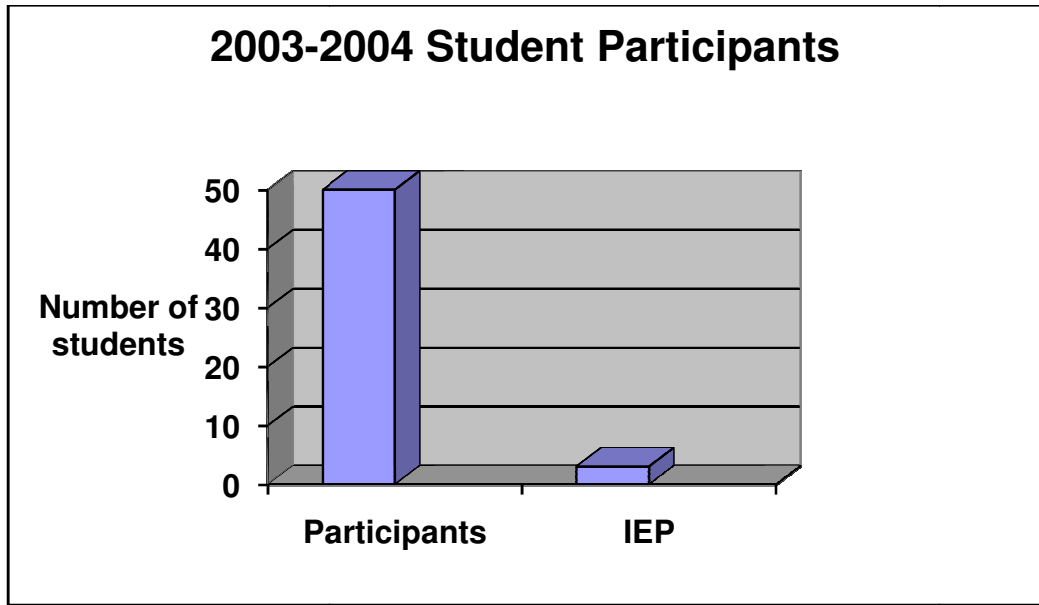


Figure 2. Student participants with IEP's for 2003-2004 school year.

As illustrated in Figure 3, during the 2004-2005 school year, 50 student subjects were selected to examine achievement levels. Of these 50 subjects, 25 were male and 25 were female. Forty-eight student subjects were Caucasians. Of the remaining two subjects, one was African American and one was Native American. During the 2004-2005 school year, the study high school did have students from other ethnic backgrounds enrolled in the junior class as illustrated in Figure 4. Through the random selection process of this study, no other groups were represented.

This study also examined the number of student subjects who were receiving additional educational services. Four of the 50 student subjects received additional educational services as documented in their IEPs as illustrated in Figure 5.

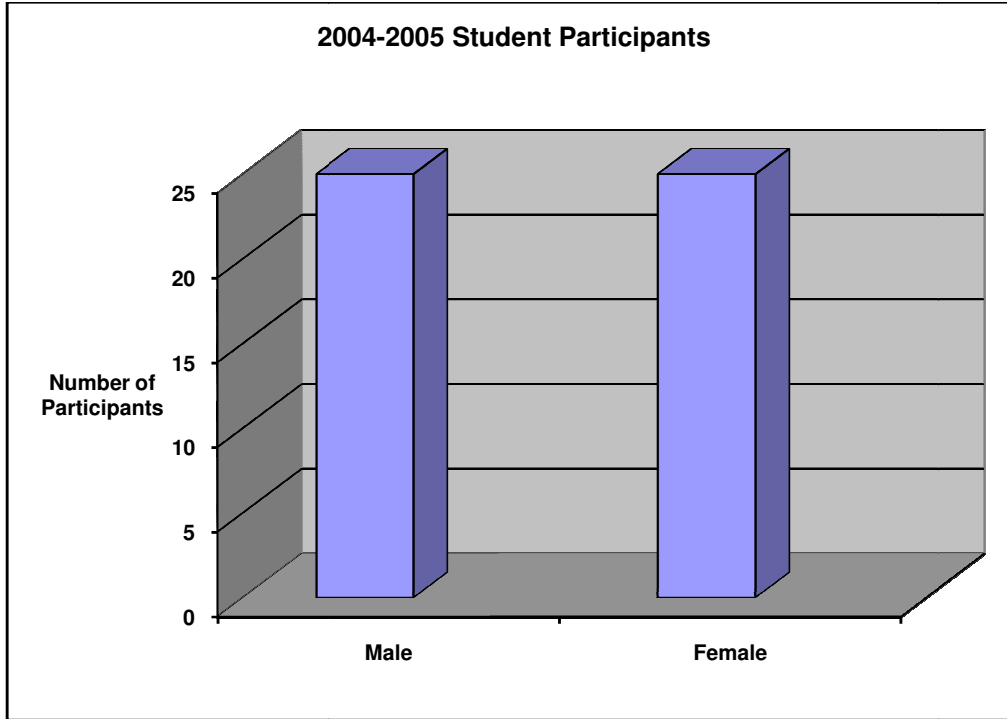


Figure 3. Gender of student participants for 2004-2005 school year.

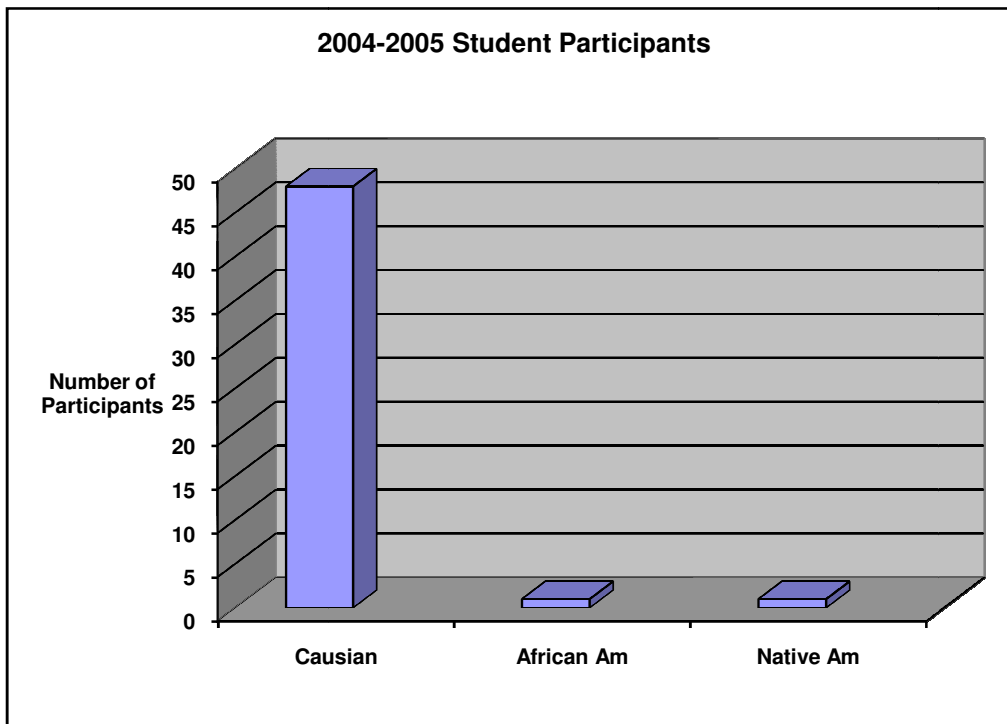


Figure 4. Ethnicity of student participants for 2004-2005 school year.

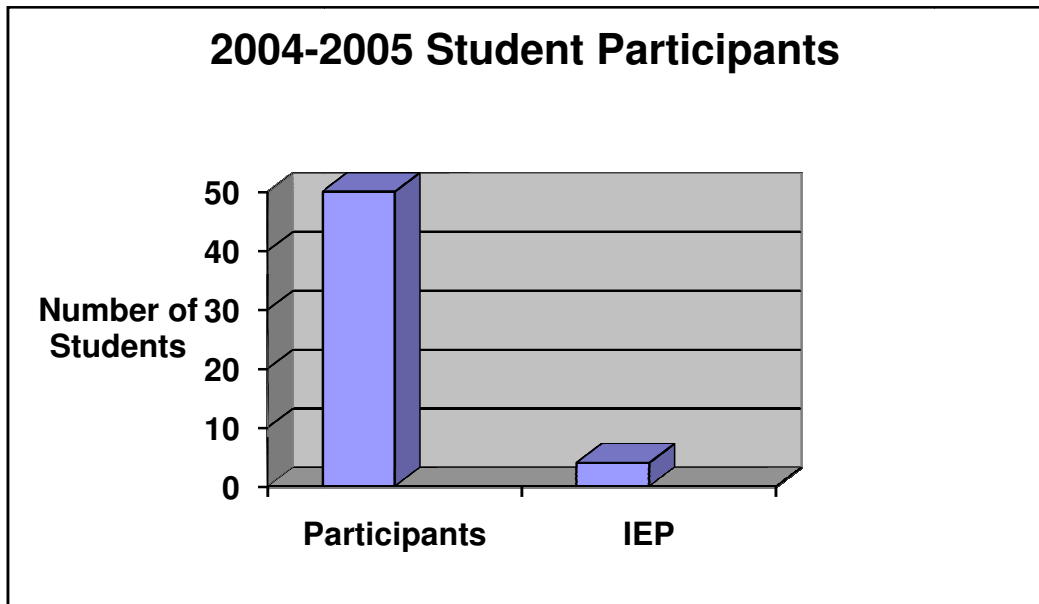


Figure 5. Student participants with IEP's for 2004-2005 school year.

Sampling Procedure

During the 2003-2004 and 2004-2005 school years, students at the study high school were placed in one of two junior English III classes. At the study high school there were only these two faculty members who taught junior English III. From these junior English III classes, 50 students were selected from each school year, totaling 100 students. Teacher A and Teacher B remained consistent during the years of the study. Both Teacher A and Teacher B taught the same district approved curriculum content to the junior students at the study high school. In the study high school teachers were required to teach the same curriculum concept at relatively the same time according to the district pacing guide. The content of the study school district's curriculum was aligned to match the Show-Me-Standards as outlined by DESE. These Show-Me-Standards were assessed by the MAP. Despite being in different sections throughout the day, the student

participants were exposed to the same curriculum for an equal amount of time during the course of the school year.

Research Setting

This study was conducted at the study high school in the study district. The high school is located approximately 20 miles south of the city of St. Louis, Missouri. The district educates over 3,000 students from three Missouri towns. The school district is comprised of many residential dwellings and supports little local industry. Because of the lack of industry, 75% of district revenue comes from state generated funds. Despite the lack of industry, district residents' income levels are equal to the middle class socioeconomic status. According to DESE's data collection, during the 2003-2004 and 2004-2005 school years, the district had a free and reduced lunch rate average of 21.3% (study school data).

Of the 3,000 students in the district, almost 1,000 of them are high school students. The ethnic backgrounds of students in the study district have seen little change in the last five years. The demographic data from the last five years has also seen little change. During the 2003-2004 and 2004-2005 school years, the study high school served 1911 students. Of the 1911 students, 1850 or 96% were Caucasian students, 3 or .015% were Indian, 14 or .07% were Hispanic, 32 or 1.7% were African American, and 12 or .06% were Asian (Study school data). Figure 6 illustrates the ethnic breakdown. It is from these high school students that student achievement data were collected.

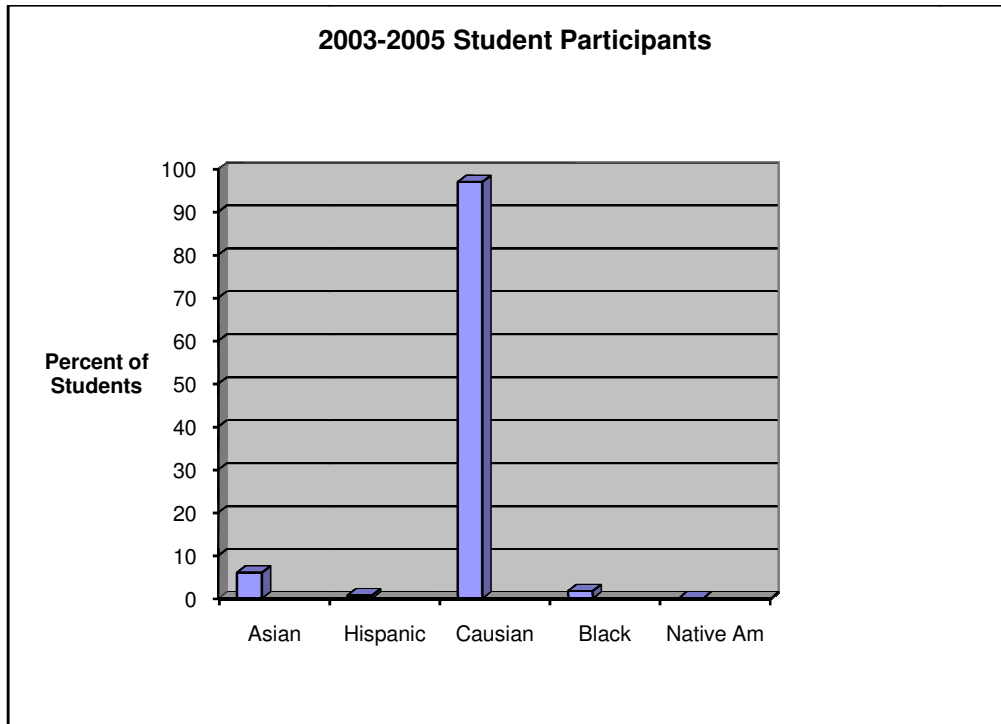


Figure 6. Ethnic breakdown of student participants 2003-2005 school years.

Over the past four years, the Communication Arts department has consisted of seven teachers certified to teach and assess Communication Arts by DESE. Of these seven teachers, two taught and assessed English III, a course required for all junior students. The junior class at the study high school averages approximately 200 students. Over the last four years, these juniors have all received classroom grades for English III and achievement levels for having taken the eleventh grade MAP. The relationship between these two data sets was examined in this study.

External Validity

The results of this study may be of use to other high schools with similar demographics to the study high school. The study high school is made up of students predominantly from the middle class social economic status.

The results of this study will be shared with the study school district school officials. Through the analysis of the results, revisions may be made in the area of classroom assessments, grading and reporting structure, and professional development. The faculty at the study high school will be able to use the data to examine their current practices to establish goals that will close any gaps in student achievement the data might expose.

Research Design

This study used a correlation analysis that examined the relationship between two sets of data. One data set consisted of achievement levels of student subjects in a junior level English class, while the other data set consisted of achievement levels of student subjects as measured by the Communication Arts MAP test. In addition, the author examined the percent of students recorded in the five achievement levels of the classroom grade compared to the percent of students recorded in the five achievement levels of the MAP using a chi-square test. The chi-square test was used because it is the most accurate measure for determining relationships for homogeneity of proportions.

During the 2003-2004 school year, the study high school had an enrollment of 197 juniors, of which 50 were selected to collect achievement levels. One data set collected was achievement levels recorded from classroom grades. Another data set collected was achievement levels recorded from the Communication Arts MAP. During the 2004-2005 school year, the study high school had an enrollment of 215 juniors. Of the 215 juniors, 50 students were randomly selected from the class roster by the researcher to collect achievement levels. Two data sets were collected to reflect achievement levels from classroom grades and the achievement levels from the MAP.

Achievement levels from classroom grades were correlated to achievement levels from the MAP test to determine if a relationship existed. In addition, the percent of students recorded at the five achievement levels of the classroom grade and MAP were examined. This process was completed for both school years.

Instrumentation

In 1993, the Missouri State Legislature passed the Outstanding Schools Act creating the MAP. The State Board of Education directed DESE to develop standards to identify the knowledge, skills, and competencies that all Missouri students should possess by the time they graduate from high school. These academic standards became known as the Show-Me-Standards. There were a total of 73 standards identified. Forty standards addressed content in all of the subject test areas and 33 were process standards. The subject areas addressed included Communication Arts, Mathematics, Science, Social Studies, Fine Arts, and Health/Physical Education. All of these standards addressed the concepts students should be able to do or perform prior to graduation (DESE, 2004).

In addition to developing standards, DESE (2004), was charged with developing an instrument to assess whether or not Missouri students were achieving or making progress toward academic standards. The instrument developed was a criterion-referenced assessment called the MAP. In 1997, DESE originally established a timetable to notify school districts of subject areas which would be assessed and when the assessment would take place. This DESE timetable also outlined the grade levels, or spans, for which the assessment would be administered. It was determined that Communication Arts, Mathematics, Science, and Social Studies would be assessed once

each at the elementary, middle, and high school levels (CTB McGraw-Hill, 1999).

Participants were assessed according to in the following schedule:

1. Grades 3 and 7: Communication Arts and Science.
2. Grades 4 and 8: Mathematics and Social Studies.
3. Grade 10: Mathematics and Science.
4. Grade 11: Communication Arts and Social Studies.

In the beginning, plans were also made to implement Fine Arts, Health, and Physical Education. In 1991, the assessment of Health/Physical education was included. However, due to cutbacks in the state budget, the assessment in this subject was not continued. The Fine Arts assessment was never administered.

The MAP test assesses students' knowledge and understanding by asking students to respond to three different types of questions. These include multiple-choice, constructed-response, and performance event questions. Multiple-choice questions have been in use by educators for a long time and are considered the standard form of assessment. The multiple-choice format presents the student with a question called the stem and provides four possible answers. The constructed-response question requires the student to supply the answer rather than select it from several choices. These types of questions ask students to supply as little as one word or as much as a couple of sentences to demonstrate understanding of the question. The constructed-response question appears in one or two forms: closed ended and open ended. The close-ended constructed-response questions require the student to provide a written answer in which there is a right or wrong answer. The open-ended constructed-response questions allow the student to provide one of many possible correct answers (CTB McGraw-Hill, 1999).

Finally, the performance event questions assess the student's ability to apply learned knowledge to solve problems. This type of question requires the students to work through or analyze a problem and provide a written answer. Together, these three types of questions make up the eleventh grade communication arts MAP (CTB McGraw-Hill, 1999).

Validity and Reliability of the MAP

CTB McGraw-Hill, in conjunction with DESE, has taken measures to ensure the MAP test is assessing the Show-Me-Standards as it was intended to do. Content experts are used to determine if assessment items are appropriate for grade level and subject areas. After the content experts determined questions are appropriate, Missouri educators are asked to review the finding of the content experts by using an item-to-standard congruence rating. Both CTB McGraw-Hill and DESE continue to examine item and score pattern analysis to ensure the results are meaningful and measuring the Show-Me-Standards (CTB McGraw-Hill, 1999).

The developers of the MAP went to great lengths to make sure the instrument was reliable. The assessment contractor CTB McGraw-Hill, developed and tested the multiple choice questions in accordance with accepted procedures and criteria intentionally aligning the questions to the Show-Me-Standards. In addition, Missouri educators wrote and reviewed alignment to the Show-Me-Standards for the constructed response and performance event questions (CTB McGraw-Hill, 1999). Despite their effort, the fact remained that test error still can come from a variety of sources such as the examiner, assessment environment, or the instrument itself. Due to test error, CTB McGraw Hill and DESE analyzed the score reliability when constructing the MAP. The MAP was

developed with both selected response and constructed response items. Because the constructed response items were judged by human readers, it was understood that some reliability may be lost when compared to those scored by a machine. Despite this loss, CTB McGraw Hill and DESE felt it was necessary to keep the constructed response items due to them being more applicable to real life situations (CTB McGraw-Hill, 1999).

Table 2 illustrates the MAP scale score reliability coefficients over a three year period.

Reliability is determined by a number ranging from zero to one. The closer the coefficient is to one the more reliable the test scores. In the case of the MAP over a three year period, the reliability coefficient was very close to one, meaning the MAP test was a reliable measure of achievement.

Table 2

Reliability Coefficients of the MAP

<u>Communication Arts</u>	<u>Year 1</u>	<u>Year 2</u>	<u>Year 3</u>
Grade 11	.939	.919	.917

Note. From Missouri Assessment Program: Guide to Test Interpretation, by CTB McGraw-Hill, 1999, Monterrey, CA: Author.

CTB MCGRAW-HILL also compared the MAP to other assessment instruments. When it was compared to other assessment instruments such as the Advanced Placement Examinations, SAT I, and the ACT Assessment, it proved to be just as reliable or, in some cases, more reliable. Table 3 illustrates the reliability information for educational assessments similar to the MAP.

Table 3

Reliability Information for Educational Assessments Similar to the MAP

<u>Assessment</u>	<u>Composite Score</u>
Advanced Placement (AP)	.85 - .96
SAT I	.91 - .93
ACT	.82 - .91

Note. From Missouri Assessment Program: Guide to Test Interpretation, by CTB McGraw-Hill, 1999, Monterrey, CA: Author.

The Classroom Grade

The classroom grade is the recorded achievement level after the culmination of many tasks the student completed during a given time frame. These tasks typically varied from one classroom to the next. The freedom that American education has allowed educators in the area of grading has been the reason for inconsistency in the overall grading system. Because of this freedom, educators usually use the classroom grade to represent several different tasks including (a) assessments (formative and summative), (b) homework, (c) additional optional tasks (extra credit), (c) attendance, (d) behavior, and (e) effort and participation. To better understand the grading structures of the two high school teachers used in this study, both the grading systems for Teacher A and Teacher B were examined. The results of this examination are illustrated in Table 4.

Table 4

Grading Structures for Teacher A and Teacher B Divided by the Percentage Assigned to each Graded Task

<u>Teacher A</u>	<u>Percentage</u>
Assessments (Formative and Summative)	70 %
Homework	20%
Additional optional task (Project)	10%
Attendance	0%
Behavior	0%
Effort/Participation	0%

<u>Teacher B</u>	<u>Percentage</u>
Assessment (Formative and Summative)	50%
Homework	40%
Additional optional tasks	10%
Attendance	0%
Behavior	0%
Effort and Participation	0%

Note. From Teacher surveys (see Appendix A).

Students' final English III grades, as defined by the above breakdown, were compared to their achievement level earned on the MAP assessment.

Data Analysis Procedure

In order to obtain the data on the randomly selected junior students at the study high school, the following procedure was used. Permission was sought from the study district officials to obtain individual classroom grades for the students randomly selected. Permission to examine the two English III teachers' classroom structure was also requested. This examination determined the structure and composition of the assigned classroom grade. Permission to review individual student MAP achievement levels was also sought from the study school district administration. After permission was granted, the students' classroom grades and MAP achievement levels were examined by the researcher by calculating and comparing outcomes by using pencil, paper and a calculator.

The classroom grades were taken from the grade books of teachers A and B. Grades were divided into achievement levels as illustrated in Table 5. The MAP achievement levels were divided by scale score as illustrated in Table 6. This table identifies the number of points a student must accumulate to earn a particular achievement level.

Table 5

Achievement Levels as Reported by the Classroom Grade

<u>Grade</u>	<u>Percentages</u>
A	97-100
A -	90-96
B+	87-89
B	84-86
B -	80-83
C+	77-79
C	74-76
C -	70-73
D+	67-69
D	64-66
D-	60-63
F	59 and below

Note. From the study high school Student Handbook, 1995, Fall.

Table 6

MAP Achievement Levels Divided into Scale Scores

<u>Achievement Level</u>	<u>Scale Score</u>
Advanced	783- 915
Proficient	738-782
Nearing Proficient	706-737
Progressing	687-705
Step 1	563-686

Note. From Missouri Assessment Program: Guide to Test Interpretation, by CTB McGraw-Hill, 1999, Monterrey, CA: Author.

Summary

The purpose of the study was to determine if there was a relationship between the classroom grade and achievement levels on the MAP test. In this chapter, details about the student participants' gender and ethnic backgrounds were discussed. The research was conducted at the study high school in the study district. The district had about 3,000 students enrolled at the time of the study with almost 1,000 students enrolled at the high school. The assessment instrument used in this study was the Communication Arts section of the MAP. The classroom grades of the student subjects were also analyzed. The structure and details pertaining to the two instruments were outlined in this chapter. Last, this chapter identified the percentages associated with particular classroom grades as well as scale scores associated with the MAP achievement levels. Chapter four reports the study results from the 2003-2004 and the 2004-2005 school years. The researcher

used a correlation analysis and a chi-square analysis to determine the significance between classroom grades and MAP achievement scores. Both of these analyses were used to determine the range of difference between the hypothesized frequencies and actual data sets.

Chapter Four - Results

Beginning with the 2002 school year, the study district had a curriculum in place that was aligned with the Missouri Show-Me-Standards. This curriculum was designed to provide the study district students exposure to the same curricular standards measured by the MAP. A standards-based curriculum provided students an equal opportunity to learn the information needed to pass a class and score in the proficient level on the MAP test. All the study district's teachers presented the curriculum concepts covered by the MAP.

The purpose of this study was to examine the relationship of the classroom grades of eleventh grade English III students to their achievement levels as evidenced by the MAP test results. The author collected data from 50 juniors for both the 2003-2004 and 2004-2005 school years. After the data had been collected, the author conducted two different analyses to determine the relationship of the achievement levels earned for the classroom grades and the MAP achievement levels to the percent of students scoring at each achievement level for both the classroom grade and the MAP. The first analysis was a chi-square test for homogeneity of proportions, and the second analysis was through calculation of a correlation coefficient.

Analysis

During the data collection process, student subjects identified by a number were listed with their final classroom grade and their MAP classification and scale score. The data collection process was the same for the student subjects selected during both the 2003-2004 and 2004-2005 school years. After all data had been collected on student participants for both the 2003-2004 and the 2004-2005 school years, the author divided student subjects into the classroom grading categories A, B, C, D, and F according to the

recorded grade earned. Student participants were also divided into MAP achievement level categories according to their recorded MAP scale score. At the study high school, the classroom grade of a B or 80% is determined to be proficient. Due to a B being proficient, the other four classroom grades were aligned to other four MAP achievement levels. The classroom and MAP achievement levels are illustrated in Table 7.

Table 7

Comparison Scale for Classroom Grades and MAP Scale Scores

Classroom					
<u>Achievement Level</u>		<u>MAP Achievement Level</u>		<u>MAP Scale Score</u>	
A	=	Advanced	=	783-915	
B	=	Proficient	=	738-782	
C	=	Nearing Proficient	=	706-737	
D	=	Progressing	=	687-705	
F	=	Step 1	=	563-686	

Note. From Missouri Assessment Program: Guide to Test Interpretation, by CTB McGraw-Hill, 1999, Monterrey, CA: Author.

2003-2004 Descriptive Statistics

After sorting the student participants into MAP achievement levels for the 2003-2004 school year, the researcher determined that zero out of 50 or 0% earned the highest achievement level of Advanced. The number of students earning the achievement level of Proficient was eight out of 50 or 16%. Twenty-five out of 50 or 50% student participants

earned the achievement level of Nearing Proficient, while 12 out of 50 or 24% of student subjects earned a MAP achievement level of Progressing. Finally, the lowest MAP achievement level, Step 1, had five out of 50 or 10% of the student participants recorded. The number of students earning these MAP achievement levels is illustrated in Figure 7.

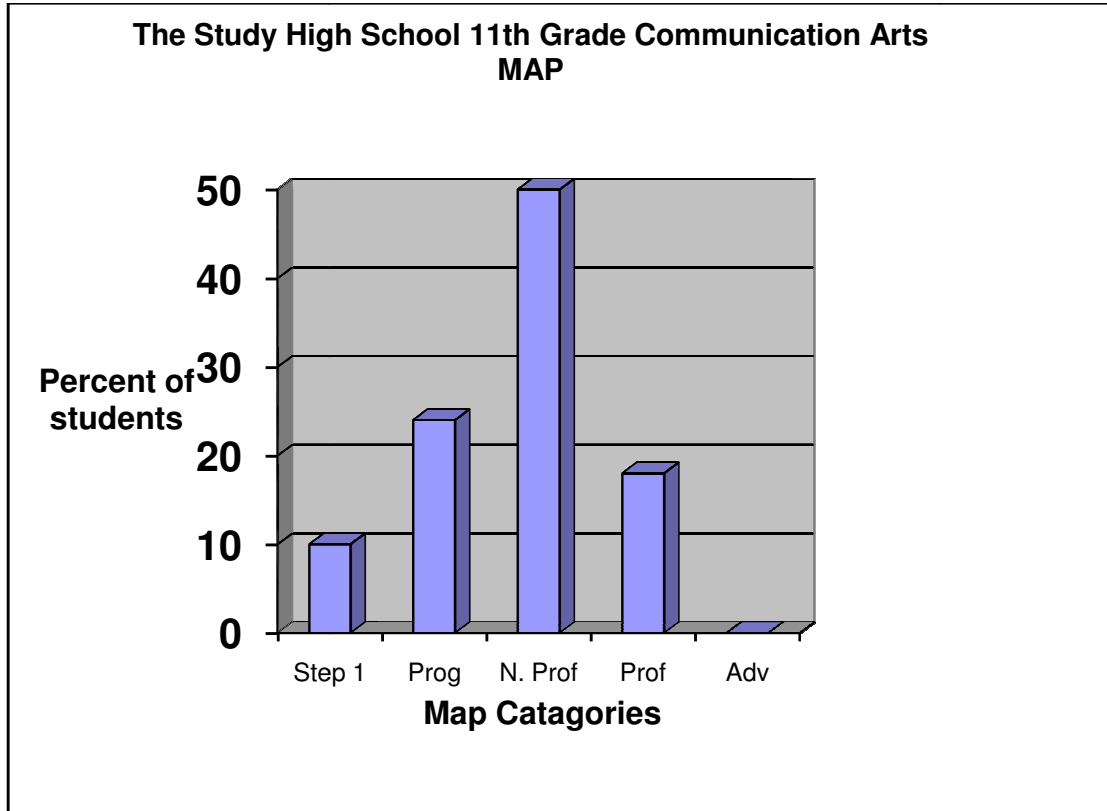


Figure 7. Student subject MAP Scores by achievement level for the 2003-2004 school year.

After sorting student subjects for the 2003-2004 school year, the researcher determined that nine of the 50 or 18% achieved the highest achievement level of an A. The number of student participants who earned the achievement level of a B was 13 out of 50 or 26%. The middle achievement level classification of C had 20 out of 50 or 40% student participants recorded, while seven out of the 50 or 14% student participants

recorded a classroom achievement level of a D. Finally, the number of students who earned the classroom achievement level of an F was one out of 50 or 2%. The number of students earning each classroom achievement level is summarized in Figure 8.

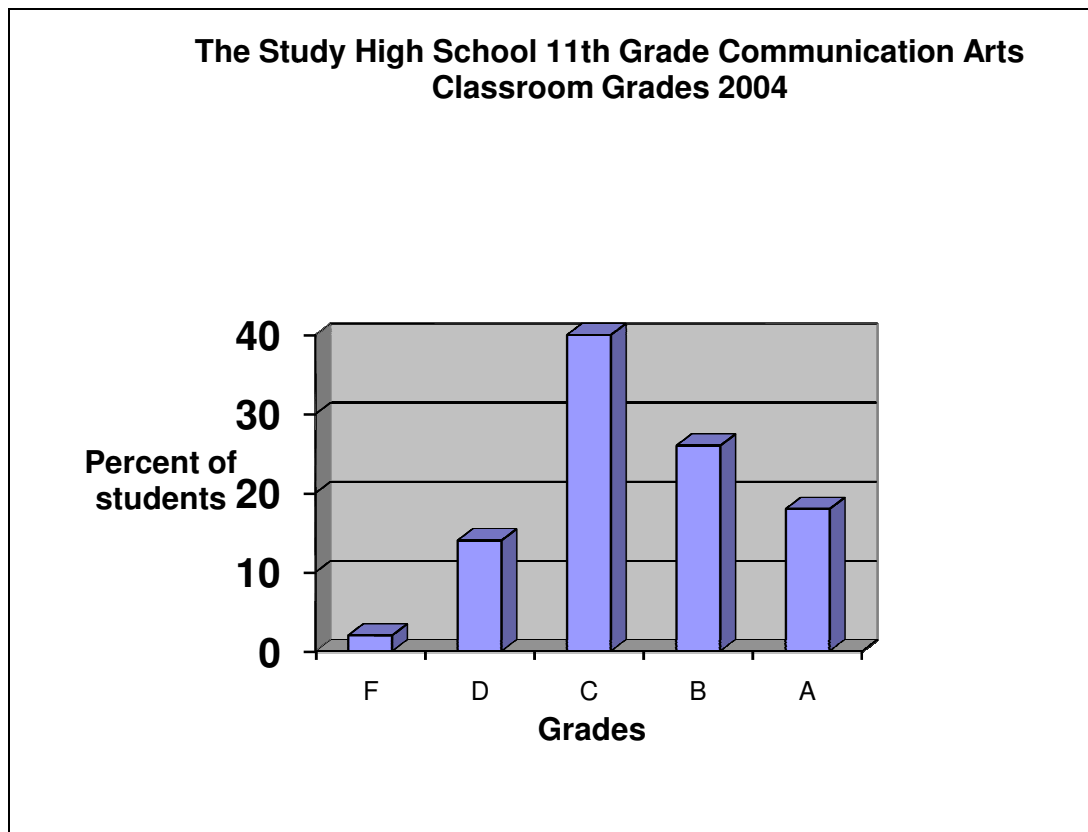


Figure 8. Student participants classroom grade for the 2003-2004 school year.

Both classroom achievement and MAP achievement were organized into five levels in which students scored. Each of the five levels represented various degrees of achievement. The highest degree of achievement for the classroom grade was an A and Advanced for the MAP. The lowest degree of achievement for the classroom grade was F and Step 1 for the MAP. During the 2003-2004 school year, 44% of the student participants scored in the top two levels of the classroom grade, A and B, while only 16%

of the student participants scored in the top two levels of MAP, Advanced and Proficient. Furthermore, while 18% of the student participants earned the highest classroom grade of A, 0% of the student participants earned the highest MAP achievement level of Advanced. Sixteen percent of participants scored at the two lowest classroom grades of D and F. Thirty-four percent of participants were recorded at the two lowest MAP achievement levels of Progressing and Step 1. Table 8 illustrates the comparison between the two data sets by showing the percent of student participants who scored at each particular achievement level.

Table 8

Percent of Students Scoring in the MAP Achievement Levels and the Classroom Achievement Levels for the 2003-2004 School Year

Percent of Recorded MAP Achievement Levels	Map Achievement Levels	Classroom Achievement Levels	Percent of Recorded Classroom Achievement Levels
0%	Advanced	A	18%
16%	Proficient	B	26%
	Nearing		
50%	Proficient	C	40%
24%	Progressing	D	14%
10%	Step 1	F	2%

Note. The percent is based on N = 50 student subjects.

2004-2005 Descriptive Statistics

After sorting the student subjects into MAP achievement levels for the 2004-2005 school year, it was determined that zero out of 50 or 0% earned the highest achievement level of Advanced. The number of students earning the achievement level of Proficient was eight out of 50 or 16%. Twenty-two out of 50 or 44% student participants earned the achievement level of Nearing Proficient, while 13 out of 50 or 26% of student participants earned a MAP achievement level of Progressing. Finally, seven out of 50 or 14% student participants were recorded in the lowest MAP achievement level, Step 1. The percent of students earning these MAP achievement levels is illustrated in Figure 9.

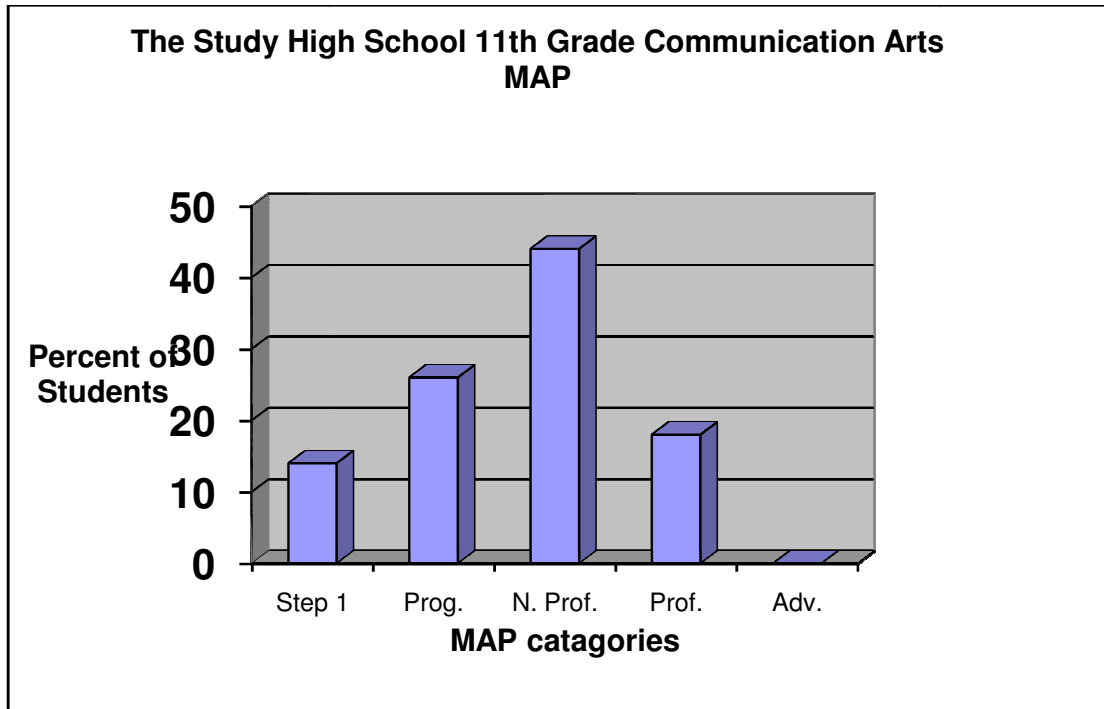


Figure 9. Student participant MAP scores by achievement levels for the 2004-2005 school year.

After sorting student participants for the 2004-2005 school year, it was determined that 15 of the 50 or 30% achieved the highest achievement level of A. The number of student participants who earned the achievement level of B was 14 out of 50 or 28%. The middle achievement level classification of C had 20 out of 50 or 40% student participants recorded, while nine out of the 50 or 18% student participants recorded a classroom achievement level of D. Finally, the number of students who earned the classroom achievement level of F was three out of 50 or 6%. The percentage of students earning each classroom achievement level is summarized in Figure 10.

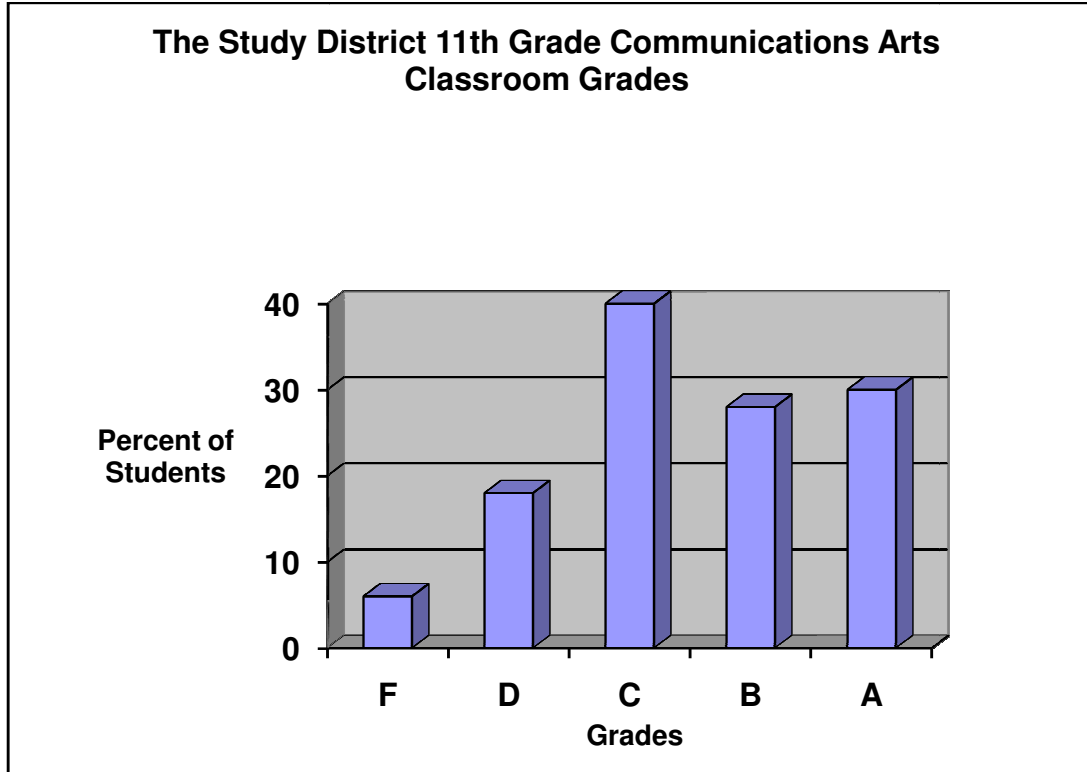


Figure 10. Student participant classroom grades for the 2004-2005 school year.

Achievement levels in which the student subjects were recorded for both the classroom achievement and the MAP achievement during the 2004-2005 school year displayed a noticeable difference in the top two levels. In the top two levels of the classroom grade (A or B), 54% of the student participants were recorded, while only 16% of the student participants scored in the top two levels of MAP (Advanced and Proficient). Furthermore, while 30% of the student participants earned the highest classroom grade of A, 0% of the student participants earned the highest MAP achievement level of Advanced. The two lowest classroom grades (D and F) had 24% of the student participants recorded. Examination of the two lowest MAP achievement

levels (Progressing and Step 1) showed that 40% of the student participants were recorded at these levels. Table 9 illustrates the comparison between the two data sets by noting the percent of student participants who scored at each particular achievement level.

Table 9

Percent of Students Scoring in the MAP Achievement Levels and the Classroom Achievement Levels for the 2004-2005 School Year

Percent of Recorded MAP Achievement Levels	Map Achievement Levels	Classroom Achievement Levels	Percent of Recorded Classroom Achievement Levels
0%	Advanced	A	30%
16%	Proficient	B	28%
	Nearing		
44%	Proficient	C	18%
26%	Progressing	D	18%
14%	Step 1	F	6%

Note. The percent is based on N = 50 student participants.

Discussion

To further examine the relationship that existed between the percent of students recorded in the achievement levels of the MAP and the classroom grade, the mean was

calculated for each. Once the mean was calculated for the classroom grade, it was then placed on the classroom grading scale. After the mean was calculated for the MAP scale score, it was then placed on the MAP index scale. Due to the classroom grading scale and MAP index scale not being divided into percentages similar to the classroom grade, it was necessary to convert the MAP index scale into percentages. The classroom grading scale was divided into thirds with the top third being the achievement level recognized with a plus, +, (example C+), the middle third recognized with a simple letter (example C), and the bottom third denoted with a minus sign, -, (example C-). To equate the structure of the classroom grading scale to the achievement levels of the MAP, the MAP index scale was divided into thirds. For example, the Nearing Proficiency level of the MAP scale score was determined to have a range of 706-737. The total range of points for the Nearing Proficiency level was a total of 31 points. The total range of points was divided by three equaling 10.3. Because the base of the Nearing Proficiency range was 706, an additional 10.3 was added to determine the bottom third. The range for the bottom third thus became 706-716.3. The middle third became 716.3-726.6, and the top third became 726.6-737. After the MAP index scale was modified, the means of both achievement levels were placed on the scales. Table 10 illustrates the means for the student participants for the 2003-2004 school year for both the MAP and classroom achievement levels.

Table 10

Mean for Classroom Grade and MAP Scale Score for the 2003-2004 School Year

Student classroom	Classroom Achievement	MAP Scale	Student MAP Scale
<u>Mean</u>	<u>Levels</u>	<u>Score</u>	<u>Mean</u>
78.7%	77-79 = C+	728-737	
	74-76 = C	716-727	
	70-73 = C-	706-715	715.86

Note. The values represent the mean achievement levels of students.

After calculating the mean for the achievement levels of the classroom grade and the MAP achievement levels, the difference existing between the two means was observed. The mean for the classroom grade of 78.7% translates into a C+ on the classroom grading scale. The mean of 715.86 for the MAP scale score translates into a C- on the classroom grading scale. If the mean MAP scale score were rounded up, it would be equivalent to a C on the grading scale. Therefore, a 5% difference existed between the classroom achievement mean and that of the MAP achievement mean. This 5% represents a half of a letter grade for the classroom achievement level.

After calculating the two means, it was determined that the mean for the 50 student subjects' classroom grade was 80.7%, while the mean for the MAP achievement was 711.92. Table 11 further illustrates where the two means fell when placed on the classroom grading and MAP scale.

Table 11

Mean for Classroom Grade and MAP Scale Score for the 2004-2005 School Year

<u>Student Classroom</u>	<u>Classroom Achievement</u>	<u>MAP Scale</u>	<u>Student MAP Scale</u>
<u>Mean</u>	<u>Levels</u>	<u>Score</u>	<u>Mean</u>
80.7%	80-83 = B-	738-747	
	77-79 = C+	728-737	
	74-76 = C	716-727	
	70-73 = C-	706-715	711.92

Note. The values represent the mean of the student achievement levels.

When the achievement levels of the classroom grade were compared to the MAP achievement levels, a difference in mean achievement levels was identified. The mean for the classroom grade of 80.7% translates into a B- on the classroom grading scale, while the mean for the MAP scale score, 711.92, translates into a C- on the classroom grading scale. Therefore, a 10% difference existed between the relationship of the classroom achievement mean and that of the MAP achievement mean. This 10% represents an entire letter grade on the classroom achievement level scale.

Correlation Analysis

A second examination of the data was conducted using a correlation analysis. This analysis was conducted to determine if a linear relationship existed between the 50 student subjects' achieved levels earned on the MAP test and their junior English III classroom grades. This portion of the study examined the relationship of the two variables by using the mathematical statistic developed by Karl Pearson known as the

Pearson Product Moment Correlation Coefficient (Runyon, Coleman, & Pittenger, 2000).

This technique is used to determine if there is a correlation between two variables. The range of the correlation coefficient is between -1.00 and +1.00. Within this range, a correlation of 0 indicates the absence of a linear relation between two variables. A positive correlation indicates there was a direct relationship between the two variables while a negative correlation indicates that an inverse relationship exists between the two variables. Correlation coefficients can be characterized as small, medium, or large as illustrated in Table 12.

Table 12

Cohen's Guidelines for Correlation Coefficients

<u>Correlation</u>	<u>Negative</u>	<u>Positive</u>
Small:	-29 to -.10	.10 to .29
Medium:	-.49 to -.30	.30 to .49
Large:	-1.00 to -.50	.50 to 1.00

Note. From *Fundamentals of Behavioral Statistics*, by R. P. Runyon, K. A. Coleman, and D. J. Pittenger, 2000, Boston: McGraw-Hill Highe Education.

Not all researchers agree on the values represented by Cohen's guidelines. For example, Runyon et al. (2000) used the coefficient of determination to interpret the magnitude of the correlation coefficient. The coefficient of determination, r^2 , is the amount of variation in the dependent variable that can be explained by knowledge of the independent variable. This study used a correlation analysis to determine the degree to

which the achievement levels of MAP scale scores and the achievement levels of the classroom grades were correlated.

For the correlation portion of the 2003-2004 analysis the hypotheses are:

H_0 : There will be no significant correlation between the Communication Arts classroom grade and Communication Arts criterion-referenced test student achievement levels as measured by MAP scores.

H_1 : There will be a significant correlation between the Communication Arts classroom grade and Communication Arts criterion-referenced test student achievement levels as measured by MAP scores.

The hypotheses for the chi-square test for homogeneity of proportions are as follows:

H_0 $p_1 = p_2 = p_3 = p_4 = p_5$. The proportion of students scoring Advanced on the Communication Arts MAP exam is the same as the proportion of students earning a grade of A in junior English III.

H_1 : $p_1 \neq p_2 \neq p_3 \neq p_4 \neq p_5$. The proportion of student scoring Advanced on the Communication Arts MAP exam is not the same as the proportion of students earning a grade of A in Junior English III.

The results of the chi-square test for homogeneity of proportions were as follows: $\chi^2(4, N = 50) = 14.73, p = .005$. The alpha level was set by the researcher as .05-.95

Since the p-value was .005, the null hypothesis was rejected in favor of the alternate hypothesis. It was concluded that the differences observed in Table 11 between proportions of students in the five categories of corresponding communication arts MAP scores and classroom grades were significantly different.

For the correlation portion of the 2004-2005 analysis the hypotheses are:

H_0 : There will be no significant correlation between the Communication Arts classroom grade and Communication Arts criterion-referenced test student achievement levels as measured by MAP scores.

H_1 : There will be a significant correlation between the Communication Arts classroom grade and Communication Arts criterion-referenced test student achievement levels as measured by MAP scores.

The hypotheses for the chi-square test for homogeneity of proportions are as follows:

H_0 : $p_1 = p_2 = p_3 = p_4 = p_5$. The proportion of students scoring Advanced on the Communication Arts MAP exam is the same as the proportion of students earning a grade of A in junior English III.

H_1 $p_1 \neq p_2 \neq p_3 \neq p_4 \neq p_5$. The proportion of student scoring Advanced on the Communication Arts MAP exam is not the same as the proportion of students earning a grade of A in Junior English III.

The results of the chi-square test for homogeneity of proportions were as follows:

$\chi^2 (4, N = 50) = 24.42, p < .001$. The alpha level was set by the researcher at .05-.95.

Since the p-value was less than .001, the null hypothesis was rejected in favor of the alternate hypothesis. It was concluded that the differences observed in Table 12 between proportions of students in the five categories of corresponding communication arts MAP scores and classroom grades are significantly different.

Deductive Conclusions

As previously identified in Tables 5 and 6 in this chapter, the author selected 50 juniors at the study high school during the 2003-2004 and 50 juniors during the 2004-2005 school years. Both classroom and MAP achievement levels were collected on the selected juniors. Using the data collected from the 2003-2004 school year, a correlation analysis was conducted. The results for the 2003-2004 school year are illustrated in Table 13. The results for the 2004-2005 school year are illustrated in Table 14.

Table 13

Correlation Data Analysis for the 2003-2004 School Year

<u>Regression Statistics</u>					
Multiple R					0.434844173
R Square					0.189089455
Adjusted R Square					0.171836039
Standard Error					10.26230239
Observations					49

	<u>Df</u>	<u>SS</u>	<u>MS</u>	<u>F</u>	<u>p-value</u>
Regression	1	1154.202033	1154.202	10.95954	0.001794518
Residual	47	4949.797967	105.3149		
Total	48	6104			

H₀: There is no correlation between X (Classroom Grade) and Y (Communication Arts MAP achievement level).

H₁: There is a correlation between X (Classroom Grade) and Y (MAP Communication Arts achievement level).

$$r(48) = .435, P = .002$$

Since $P < .05$, the null hypothesis was rejected; it was concluded that a significant positive correlation existed between achievement on the MAP and the achievement of the classroom grade.

Table 14

Correlation Data Analysis for the 2004-2005 School Year

<u>Regression Statistics</u>					
Multiple R			0.470473		
R Square			0.221345		
Adjusted R Square			0.204778		
Standard Error			11.54957		
Observations			49		

	<u>Df</u>	<u>SS</u>	<u>MS</u>	<u>F</u>	<u>p – value</u>
Regression	1	1782.187	1782.187	13.36047	0.000646539
Residual	47	6269.446	133.3925		
Total	48	8051.633			

H₀: There is no correlation between X (Classroom Grade) and Y (MAP Communication Arts achievement level).

H₁: There is a correlation between X (Classroom Grade) and Y (MAP achievement level).

$$r(48) = .47, P = .0001$$

As with the previous year, since $P < .05$, the null hypothesis was rejected, and it was concluded that there was a significant positive correlation between achievement on the Communication Arts MAP and the classroom grade.

Comparison of Data Analyses

After analyzing the data using the chi square test and a correlation analysis, it was determined that the two presented opposing points of view. The chi square analysis revealed there was a difference between the achievement levels of the classroom grade and the MAP. During the years the data were examined, the null hypotheses were rejected. They were rejected due to the proportion of students in the categories of the MAP and classroom grades not being equal. This indicated a lack of consistency between MAP and classroom grades. Conversely, the correlation analysis revealed a positive relationship existed between achievement in the classroom and achievement on the MAP. As MAP scores increased, so did classroom grades. This suggests consistency between MAP and classroom grades.

The correlation between the student subjects' classroom grades and MAP achievement levels were further illustrated by using a scatter plot. Figure 11 demonstrates the linear relationship between the two variables for the student subjects during the 2003-2004 school year. From the scatter plot, it was determined that a positive relationship existed between the two variables. Figure 12 demonstrates the linear relationship between the two variables for the student subjects during the 2004-2005 school year. Just as in the

previous year's scatter plot, it was determined that a positive relationship existed between the two variables.

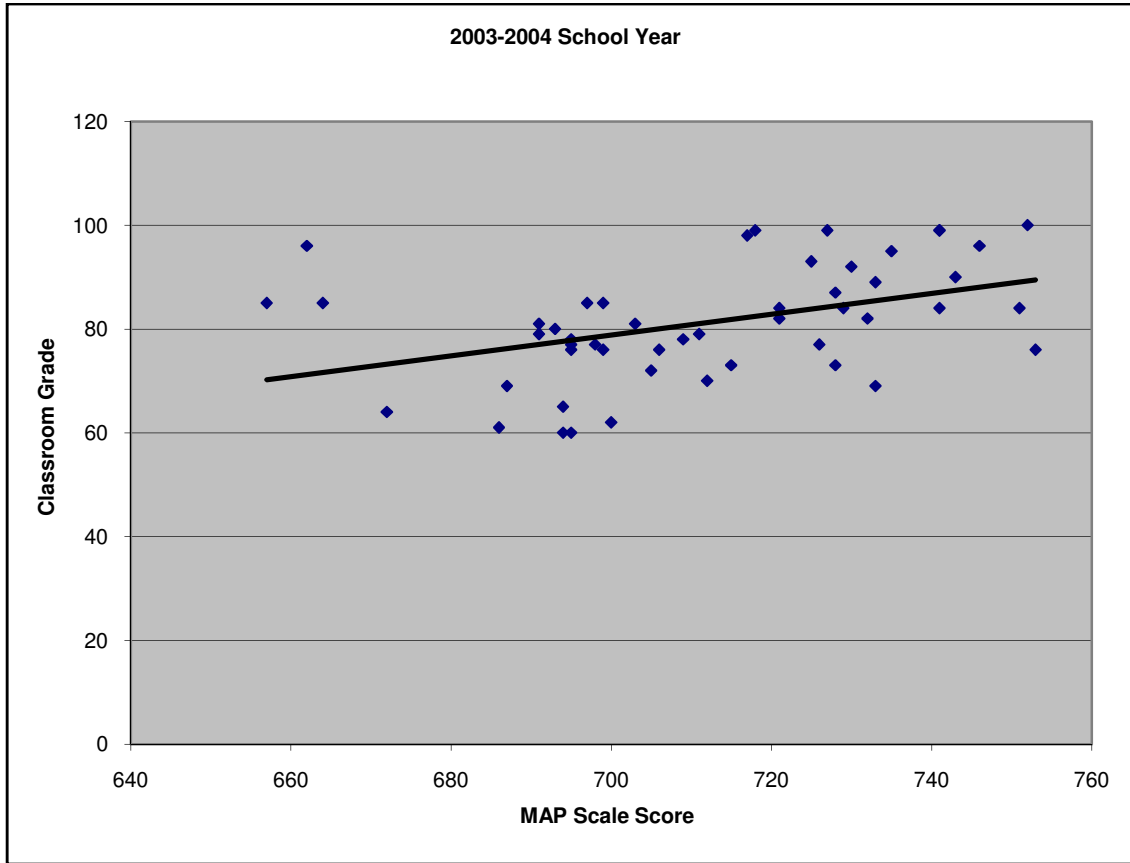


Figure 11. Scatter plot correlating MAP scale scores to classroom grades for the 2003-2004 school year.

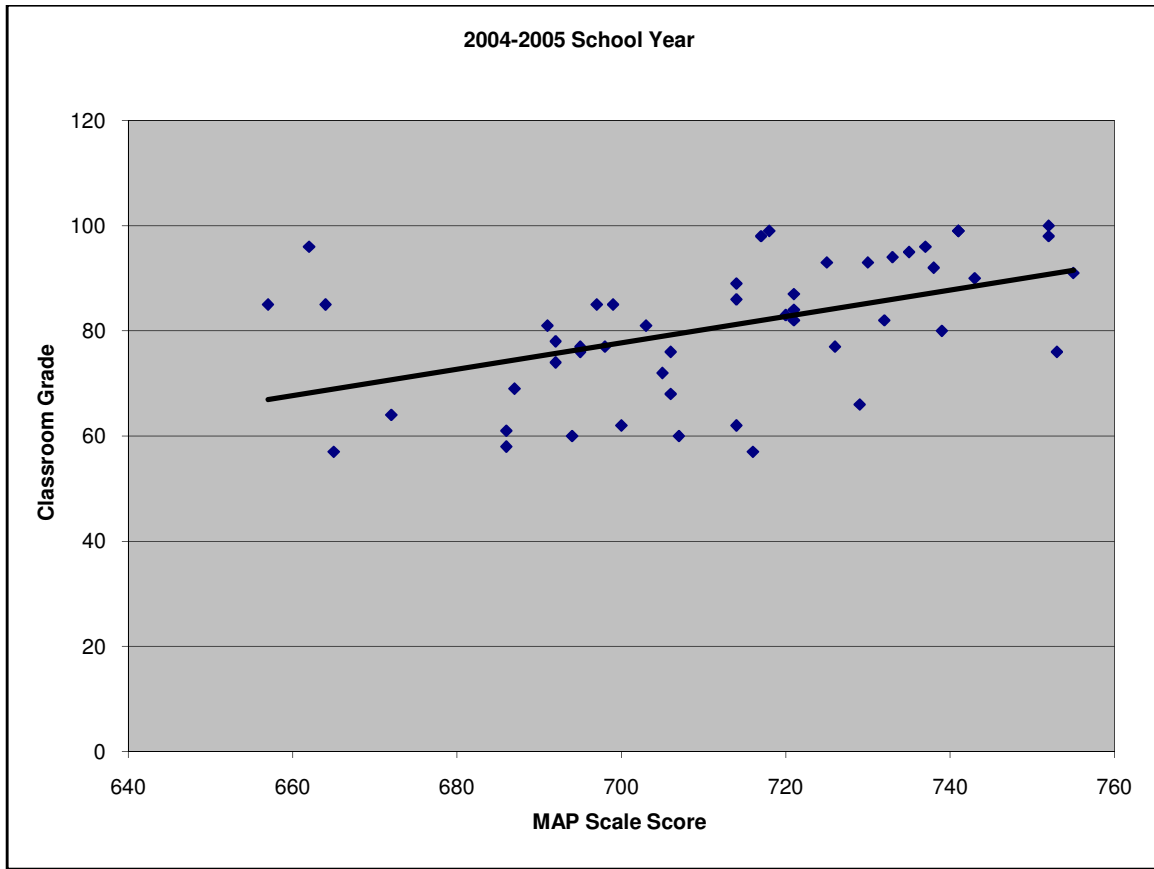


Figure 12. Scatter plot correlating MAP scale scores to classroom grades for the 2004-2005 school year.

Summary

To determine if a relationship existed between classroom grades and achievement on the MAP test, data were collected from the records of the study high school junior participating. Data for 50 student subjects were collected from the 2003-2004 school year, and 50 student subjects from the 2004-2005 school year. The data were analyzed using both a chi square test and a correlation analysis. In conducting the chi-square analysis, the data were organized into two different data sets. The first data set was organized by classroom achievement levels, while the second data set was organized by

MAP achievement levels. After the data had been organized into the two different categories, they were analyzed to determine if a relationship existed between the two by examining the number of student subjects scoring at each achievement level.

The researcher determined that there was a difference of at least 5% of a letter grade in the data examined during the 2003-2004 school year and a difference of at least 10% of a letter grade in the data examined during the 2004-2005 school year. Based on the comparison of the study high school data for the 2003-2004 and 2004-2005 school years, there were differences in achievement between classroom grade and achievement measured by the MAP test.

A second comparison was made using a correlation analysis. The results were that the null hypotheses were rejected due to significant correlations between MAP scale scores and achievement levels of the classroom grade. Despite being rejected, examination of the r^2 illustrated that during the 2003-2004 school year, $r^2 = 0.19$, meaning 19% of variability in the classroom grade can be explained by knowledge of the MAP scale score. Therefore, 81% variability in the classroom grade remains unexplainable. During the 2004-2005 school year, $r^2 = 0.22$, meaning 22% of variability in the classroom grade can be explained by knowledge of the MAP scale score. Therefore, 78% variability in the classroom grade remains unexplainable. Chapter five discusses the results and makes recommendations based on the finding of the study.

Chapter Five - Discussion

The sample setting was examined to analyze the percentages of student participants recorded in the five classroom achievement levels to the percent of student subjects recorded in the five Communication Arts MAP achievement levels using the chi-square test. In addition, a correlation analysis was conducted between student classroom achievement and the Communication Art MAP scale scores. After reviewing the data, no relationship was suggested between the achievement level earned in the classroom and on the MAP. However, the correlation analysis revealed there was a correlation that did exist between the achievement levels of the MAP and the classroom grades.

Analyzing the data using the correlation analysis for the 2003-2004 school year revealed an $r^2 = .19$, meaning a 19% of variability in the classroom grade can be explained by knowledge of the MAP scale score. However, calculating $1 - r^2$, it is indicated that 81% of the variance between the two achievement levels could not be described or explained. The correlation analysis for the 2004-2005 school year revealed an $r^2 = .22$, meaning that a 22% of variability in the classroom grade can be explained by knowledge of the MAP scale score. However, calculating $1 - r^2$, it is indicated that although there was 22% variability in the classroom grade that can be explained by knowledge of the MAP scale score, 78% of the variance between the two achievement levels could not be described or explained.

Due to these unexplained factors, perhaps there is a need for revision of the grading structure to be able to report more consistent and accurate academic achievement. Despite the conflicting findings, the traditional grading structure that prevails in the American education system may be in need of revision due to possible

grade inflation or deflation by teachers. A reason for the revision is to more accurately communicate the actual student learning or achievement occurring within classrooms. NCLB required educational leaders to address the areas of curriculum, instruction, and assessment. Beyond summative assessment, the issue of grading and reporting has yet to be addressed. Because of the failure to address this area, American schools continue to allow educators to grade and report on student achievement with inaccurate and inconsistent methods. Until the grading and reporting system is revised, American students will continue to be subjected to unequal playing fields. How one English III teacher grades and reports achievement will be different than another English III teacher. If American schools expect teachers to teach the same curriculum, practice the same instructional strategies, and use the same common assessments, why would they not be expected to grade and report achievement using the same consistent assessment system?

Educational leaders must understand the need for revising the grading and reporting structure. In addition to understanding the need for structure revision, educational leaders need to know how to implement this revision in their schools. To help educational leaders identify the steps necessary to revise the grading and reporting system in their school or district, the leadership responsibilities suggested by researchers Marzano et al. (2005) of McREL were applied to the steps. The Marzano et al. (2005) research on Balanced Leadership outlined 21 leadership responsibilities related to student achievement. The results of this study were used to create a leadership framework supported by 30 years of research data. This leadership framework is predicated on the notion that effective leadership means more than simply knowing what to do; it is knowing when, how, and why to do it. According to Marzano et al. (2005), effective

leaders must learn to compromise the existing concepts of culture, values, and norms of schools with the needed changes to move forward. These 21 responsibilities represent a balance of knowledge and skills that leaders must exhibit to positively impact student achievement. As referenced in Chapter Two, the higher the r score, the more significant the leadership responsibility. Table 15 represents the leadership responsibilities and their significance as measured by Average r scores in a contrived organizing construct.

Table 15

Leadership Responsibilities and Their Statistical Significance as Measured by Average r

Scores: An Organizing Construct to Create New Meaning.

Responsibility	Average r
<u>Change Agent</u>	.25
Ideals/Beliefs	.22
<u>Focus</u>	.24
Optimizer	.20
Situational Awareness	.33
<u>Communication</u>	.23
Intellectual Stimulation	.24
Input	.25
Flexibility	.28
Monitoring/Evaluating	.27
Outreach	.27
Relationships	.18
Visibility	.20
<u>Culture</u>	.25
Discipline	.27
Affirmation	.19
Order	.25
Contingent Rewards	.24
<u>Curriculum/Instruction/Assessment</u>	

Knowledge of Curriculum, Instruction, and Assessment	.25
Involvement in Curriculum, Instruction, and Assessment	.20
Resources	.25

Note. From School Leadership that Works: From Research to Results (pp. 42-43), by R. J. Marzano, T. Waters, and B. A. McNulty, 2005, Alexandria, VA: Association for Supervision and Curriculum Development. Adapted by the author categorizing responsibilities into strands.

As discussed in Chapter Two, Marzano et al. (2005) conducted a meta-analysis on the 21 leadership responsibilities that impact student achievement. These leadership responsibilities served as a guideline for how the results of this study were applied to revise classroom grading and reporting structures to more accurately reflect student achievement. Marzano et al. (2005) identified the 21 responsibilities as isolated acts and listed them in alphabetical order. There were similarities among the responsibilities; therefore, the 21 responsibilities were grouped by the researcher into four different strands using the organizing construct illustrated in Table 17. The construct represents the responsibilities in an organized fashion for ease of understanding. Three of the 21 leadership responsibilities – Change Agent, Communication, and Culture – seemed to be natural strand titles under which the majority of the remaining responsibilities could be organized. However, one strand title seemed to be missing, it was contrived – Curriculum/Instruction/Assessment – under which the remaining 2 leadership responsibilities were logically placed.

Despite the Average r score defining the significance of each of the responsibilities, the responsibilities grouped together in strands had a wide range. In the first strand, Change Agent, the responsibility of Situational Awareness had the highest r score of .33. This indicates that educational leaders who exhibit this responsibility will have more of an impact on student achievement. In the second strand, Communication, the responsibility of Flexibility had the highest impact at .28. Also, noted within this strand, Communication had three of the five most significant scores. In the third strand, Culture, Discipline was identified to have the greatest impact on student achievement at .27. This was the same r score as Monitoring/Evaluating and Outreach in the Communication strand. The fourth strand of Curriculum/Instruction/Assessment did not have a responsibility identified as high as the others. One possible explanation for this could be that in most cases, curriculum, instruction, and assessment are usually delegated to other persons by the school leader. This would result in many persons being involved in this area other than just the school leader. Many of the other leadership responsibilities are acts generally carried out by one person.

Implications and Recommendations for Effective Schools

Having organized the 21 leadership responsibilities in a fashion easier to understand, each responsibility will be discussed as it relates to introducing and maintaining a change to the grading and reporting structure. When implementing a new grading and reporting structure, the school leader will be challenged to carry out each of the responsibilities. Despite a large number of responsibilities, each must be present to effectively implement the suggested change. Due to the fact that only one subject area

was analyzed, additional subject areas could be analyzed to determine the significance between the classroom grade and achievement on a standardized test.

Change Agent. The leadership responsibilities associated with being a Change Agents are Ideals and Beliefs, Focus, Optimizer, Situational Awareness, and that the leadership responsibilities identified as Change Agents refer to a school leader's disposition to challenge the status quo (Marzano et al., 2005). When enacting any challenge or change, a school leader assesses the need for change by examining the current reality of the school.

After schools or districts have already undergone the mandated revisions to curriculum, instruction, and assessment, the next step should be revising the grading structure. Schools that allow inconsistent grading and reporting structures are not doing what is best for their students. Educational leaders are responsible for providing consistent grading and reporting structures for both the students and the staff. If a school leader allows inconsistent grading practices, both students and staff will supplement with what they know to be right based on their individual beliefs. While implementing a new grading and reporting structure to equalize the playing field, the school leader could temporarily upset the school's equilibrium. In doing so, this leader must be willing to lead the change initiative with uncertain outcomes while considering this study's practice recommendations.

According to Marzano et al. (2005), to be an effective Change Agent, a school leader must possess well-defined ideals and beliefs about the school's mission and what the school can accomplish. When implementing this change, the school leader must communicate and model behavior consistent with the school's vision and values with the

staff and students. By communicating and modeling, the school leader can help the staff members and students understand why decisions are being made and why a revision of the current structure could be necessary.

As stated in Chapter One of this study, many changes have been made in American education. Very few of the changes implemented in recent history have proved successful. Because of these failures, experienced educators have become skeptical of implementing new programs or creating new structures. In revising the grading and reporting structure, the school leader must make the effort to ensure the staff is focused on the purpose of the change. Marzano et al. (2005) stated the purpose can be defined by creating achievable time-bound goals. The effective school leader will continually monitor whether or not these goals are being met in a timely manner. A Change Agent is one who can provide Focus to students and staff during the revision process.

In implementing a new grading structure, the school leader must be an Optimizer. This is someone who remains positive as the driving force behind the creation of the new structure (Marzano et al., 2005). The leader must remain supportive when this revision process presents challenges. The Optimizer announces to the staff that revising the grading and reporting structure will be challenging, but the support will be there until the process is complete. Marzano et al. (2005) further elaborated, because most changes disrupt the equilibrium of a school's culture, the school leader needs to display the leadership responsibility of Situational Awareness. This allows the school leader to anticipate potential problems that may arise when implementing the change process.

Communication. This strand includes the leadership responsibilities of Intellectual Stimulation, Input, Flexibility, Monitoring/Evaluating, Outreach, Relationships, and

Visibility. Through the McREL research, it was determined that Communication had a small correlation with regard to its impact on student achievement according to Pearson Product Moment Correlation Coefficient. Marzano et al. (2005) described communication as “when a school leader establishes strong lines of communication with and between teachers and students” (p. 46). To create open and effective lines of communication, a school leader must be accessible. When initiating change to a well-established practice, such as the grading and reporting structure, school leaders must create stakeholder ownership by involving staff or teachers in the decision-making process. Important to the process is communicating the differences that exist between the proposed and current grading structures with the school building or district. Involving teachers in the discussion and research to examine other options available that could serve as better grading and reporting structures for student achievement is important in creating teacher ownership of the new change.

The school leader who ensures the school faculty is using the latest research and theory is one who provides Intellectual Stimulation as a leadership responsibility. After the research has been done, the school leader needs to bring about consensus on a structure that will meet the needs of their students. After the research is complete, the school leader (with the help of the staff) should bring about a consensus on a structure allowing the staff opportunities to have input and offer other points of view.

After a structure has been agreed upon, the school leader should then communicate to all stakeholders the new grading and reporting structure and the research used in its development. In addition, the new structure should be thoroughly explained to the school community. Like teachers, parents only understand the traditional grading and

reporting structures, so further explanation on how it will provide them with more accurate feedback will be necessary.

Upon implementation of a new grading and reporting structure, continual Monitoring and Evaluating will help to ensure its success. Because it is a new structure, some may struggle with its intended purpose. Some may have the tendency to resort back to the system they already know and feel most comfortable using. A monitoring system should be in place to ensure consistency in the implementation by the faculty and accuracy of feedback about student achievement.

Monitoring and Evaluating the new structure, the school leader should engage in Outreach. Marzano et al. (2005) described outreach as “being when the leader is an advocate and a spokesperson for all the school’s stakeholders” (p. 58). As a spokesperson, the school leader can better explain the purpose of the new grading structure and communicate that the new structure is working on providing more accurate feedback. This communication can be sent out in monthly memos to both community and school stakeholders.

Culture. The leadership responsibilities of Discipline, Affirmation, Order, and Contingent Rewards seemed to fit under the culture strand. All schools have developed a culture based on their mission, vision, and values. It is through this culture that schools come to develop goals that eventually impact student achievement. According to Marzano et al. (2005), effective school leaders build cultures that have a positive influence on teachers as well as students. In developing such a culture, school leaders should promote an understanding of the purpose of the change. This purpose would be centered on doing what is right for the students. The traditional grading and reporting

structure used throughout American education was developed based on what educators experienced as students themselves or what they feel is fair for the student. Because of the numerous structures that educators have experienced, their different operating definitions of fairness and inconsistency in the grading structure have been accepted.

The effective school leader should develop a culture where all staff members share the same vision and demonstrate the values necessary to accurately assess student learning over time. Just as collaboration of curriculum and instructional practice has already taken place, it seems to be the time for collaboration on grading structures.

The first leadership responsibility related to Culture is Discipline. Marzano et al. (2005) believed discipline involves the important task of protecting teachers from internal and external factors that would distract them from their instructional time or focus. This responsibility is critical when attempting to create buy-in from resistant staff members and misinformed community members. The school leader needs to protect and support those who are committed to developing and implementing a new grading structure. One way that this may be made easier for the school leader is by creating Order. The second leadership responsibility related to culture, Marzano et al. (2005) reported that in creating Order, “The school leader establishes standard operating principles and routines through collaboration which can ensure that everyone is on the same page when implementing and executing the new grading structure” (p. 57). In addition, it will help to identify those who are not on board or having trouble with what is being implemented.

In a school culture that is going through a change process, it is always necessary to Affirm and recognize Contingent Rewards to those who are doing what is being asked. After the new grading structure is in place, a school leader might point out someone who

has done well in adapting to this new structure. The school leader could praise staff members and students who have displayed more accurate growth through the new grading structure.

Curriculum, instruction, and assessment. The specific leadership responsibilities of (a) Knowledge of Curriculum, Instruction, and Assessment; (b) Involvement in Curriculum, Instruction, and Assessment; and (c) Resources will be discussed.

All school leaders should have knowledge of and be involved in the development, implementation, monitoring, and evaluating of curriculum, instruction, and assessment. In order to be involved, school leaders must be aware of the most current research and best practices used in curriculum, instruction, and assessment. School Leaders need to be hands-on when working with staff members in these areas. School leaders who are effective instructional leaders meet with their staff on a regular basis to discuss issues and the research that surrounds these areas. This can be done by conducting book studies or examining what has been working with other successful schools. Before changing the grading structure, a comprehensive analysis of the relevant research needs to be conducted by the staff as led by the school leader.

When initiating change to an existing structure, school leaders must provide the necessary Resources to support the change. After a new grading structure has been proposed, the school leader must ask the staff what type of resources they will need in order to successfully implement the new structure. Attention to professional development is mandatory. Staff members must to be trained on how to adjust and transition to the new grading system from the old. In addition, due to the differences in the new and existing structure, it will undoubtedly have an impact on the classroom curriculum and

instruction implementation and delivery. The greatest resource that must be provided to the staff is time. Time must be allocated to each staff member so that he or she can collaborate with others on the successes and challenges faced in the implementation of the new structure.

Relating Finding to O'Connor's (2002) Model

Although this study did not find a significant relationship between the classroom grade and the MAP scale score using a correlation analysis, it did find a relationship between the percent of students scoring in the top categories of the grading scale and the top categories in the MAP scale. Because of this second finding, there is a need to examine grading and reporting structures currently across America. The grading structure that is recommended by O'Connor (2002) is the standards-based guidelines designed to support learning and to encourage student success. These guidelines are identified in Table 16. Recall in Chapter Two that the purpose of grades is to inform students, parents, and others on the achievement of individual students (Airasian, 1994). In order for teachers to be able to accomplish this task, they must be clear on what they are grading. Traditional grading involves assigning a single number or letter to multiple curriculum concepts. In standards-based grading, students receive grades or marks based on the assessment of only one curriculum concept. By assessing only one concept at a time, all stakeholders can receive more specific feedback. In some cases, an assessment could measure a clustering of curriculum concepts. By taking this approach, it is important to understand that the number of curriculum concepts clustered must be kept to a minimum.

Table 16

O'Connor's Guidelines for Grading in Standards-Based Systems

1. Relate grading procedures to the intended learning goals	5. Grade in pencil – keep records so they can be updated easily
2. Use criterion-referenced performance standards as reference points to determine grades	6. Crunch numbers carefully – if at all
3. Limit the valued attributes included in grades to individual achievement	7. Use quality assessment(s) and properly recorded evidence of achievement
4. Sample student performance – do not include all scores in grades	8. Discuss and involve students in assessment, including grading, throughout the teaching/learning process

Note. From *How to Grade for Learning: Linking Grades to Standards* (pp. 243-244), by K. O'Connor, 2002, Arlington Heights, IL: Corwin Press.

O'Connor's Guideline #1: Relate grading procedures to the intended learning goals. Classroom teachers generally teach students specific curriculum concepts. These concepts can come in many forms whether they are concepts documented in a districts' curriculum or concepts documented from a textbook or classroom resource. When NCLB was passed, many states began developing and implementing statewide curriculum concepts to be implemented in classrooms. In the state of Missouri, DESE developed the Grade Level Expectations (GLE's). The GLE's consisted of Strands, Big Ideas, and Concepts for the subject areas of Communication Arts, Mathematics, Science, and Social

Studies for every grade level. The GLE's were developed with the idea that they would be the intended learning goals for Missouri classrooms. During the course of this study, the GLE's had just been approved by the State Board of Education, and districts were expected to begin implementing them into the curriculum. Teachers A and B in this study may not have implemented the GLE's in the manner DESE intended. In addition, the study district did not have a monitoring device in place to determine if in fact all the GLE's were implemented and at the knowledge level that would be assessed on the MAP. Furthermore, when a teacher implements a concept, the teacher must have a clear understanding in order to successfully implement it. Again, there is no evidence to support that either teacher A or B had a clear understanding of the concepts outlined by the GLE's.

Because the teachers in this study may not have implemented the GLE's in the manner that was intended, student classroom grades may have been distorted when compared to the achievement levels of the MAP. This researcher determined this through conversations with the two English III teachers at the study high school. The results produced by the chi-square test could support this possibility. After intended learning targets have been identified and understood, the classroom grade should reflect a direct correlation to each of the learning targets. Both teachers' (A and B) grade reporting did no such thing. Their classroom grades were representations of a combination of multiple items that are outlined in Table 4.

During the reporting process, as it relates to the intended learning target, it will be necessary to change the look of the school's report card. Just as individual or clustered curriculum concepts are graded and recorded in the teachers' grade books, they should be

reported to stakeholders in the same format. Traditional report cards simply state the subject area and then assign a letter grade to communicate achievement of multiple curriculum concepts. Such was the case with the study districts high school reporting process. When reporting achievement through standards-based grading, a grade or rating is assigned to multiple standards within the subject area. The student still may receive an overall grade for the subject area, but this grade will reflect achievement summary of achievement of multiple standards.

O'Connor's Guideline #2: Use criterion-referenced performance standards as reference points to determine grades. Guideline number two illustrates the need to use criterion-referenced performance standards to determine student grades. Criterion-referenced assessments allow students to measure their understanding over time against themselves rather than against another student's understanding of the standards.

Traditional practice consisted of assessments where students were assigned a grade based on the performance of a group, such as grading on a bell curve. By grading on a bell curve, only a certain number of students could achieve at a specific achievement level. For example, a student might have earned a score of a 92%, which would be an A grade on most traditional grading scales, but would receive a B grade due to a high number of students scoring above 92%. This method makes learning a highly competitive activity in which students compete against one another for the few scarce rewards (high grades) distributed by the teacher. Guskey (1996) pointed out that "learning then 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 using criterion-referenced performance standards, students are graded against themselves. The focus of these types of criterion-

referenced assessments is to demonstrate achievement, growth, and progress toward individual curriculum concepts.

Although it was uncertain if either teacher A or B implemented a norm-referenced grading system by using a bell curve, it would have impacted the distribution of the classroom grades. During the 2003-2004 school year, the data collected on classroom grades could support the possibility that perhaps a norm-referenced scoring system was used. However, during the 2004-2005 school year, the data collected on classroom grades does not support this based on interviews conducted by the researcher.

When implementing performance standards, performance levels also must be established. As indicated in Chapter Two, performance levels have been around for some time and can come in many forms. The study district high school had performance levels established prior to the study being conducted. These performance levels were identified in Table 4.

O'Connor's Guideline #3: Limit the valued attributes included in grades to individual achievement. The third guideline emphasizes limiting the valued attributes that usually are represented in reporting of student achievement. O'Connor (2002) stated, "Grades should be based on achievement, i.e. demonstration of knowledge and skill components of the standards" (p. 89). As mentioned in Chapter Two, many teachers feel compelled to include multiple variables in the reporting of the students' grades (Bailey and McTighe, 1996). These variables often include many non-academic tasks such as attendance, effort, and participation. Because non-academic tasks are not an accurate reporting measure of student achievement, they should not be included in the students' grades. According to O'Connor (2002), "Non-academic tasks such as effort,

participation, attitude and other behaviors could be reported separately. Grades should be based on individual achievement” (p. 104).

For schools to more accurately report on individual achievement, they must modify report cards. This modification could be accomplished by separating the academic achievement from the non-academic task. Just as it is important to report on individual achievement, it is equally important to report on non-academic competencies such as attitude, learning skills, and effort. By reporting on these competencies separately, more effective communication could be given to third parties in regard to a particular students’ overall achievement. Regarding teacher A and B, their grading composition included 10% that was categorized as other optional tasks. These other optional tasks could have taken the form of participation or group work. If the student performed well in the academic areas but poorly in the non-academic areas, their grade could have been penalized. This type of grading system does not allow for accurate communication and reporting of student learning.

O’Connor’s Guideline # 4: Sample student performance – do not include all scores in grades. The traditional American grading practice has been one in which teachers assign grades or marks for every task assigned to the student. In the standards-based grading structure, teachers are encouraged to sample student performance over a period of time. Teachers’ may not necessarily assign a grade or mark to every task assigned. By implementing this approach, teachers may use more formative assessments. Recall that formative assessments are snap shots of student learning as they practice new concepts. Because formative assessments are considered practice, they could not be graded. Instead, the teacher should provide descriptive feedback to students on their

progress toward reaching the desired learning goal. Descriptive feedback has proven to raise student achievement levels far more than teacher assigned grades. During a study by Page (1958), 74 secondary teachers administered tests to students in their classes. After the test was administered, they were divided into three groups. The first group's test was scored by using only a numerical score and a letter grade. The second group was scored with a numerical and letter grade and a standard comment. The third group was scored using a numerical and letter grade along with individualized comments. To evaluate the different types of feedback affects, results from the next test administered to the three groups were examined. It was found that the second group with the numerical and letter grade along with the standard comment did significantly better than those in the first group who just received a numerical and letter grade. However, the students in the third group who received individual comments did even better than the students in the second group. This study illustrated that giving individual descriptive feedback had a greater impact on student achievement (Page, p. 176). Additional studies have been conducted more recently highlighting the same results. Butler (1988) reported students who receive only numerical or letter grades demonstrate no positive growth. Further, students who receive individual descriptive feedback see as much as a 30% increase in student achievement.

An additional aspect of O'Connor's Guideline #4 focuses on sampling student performance on the summative assessment. Traditional summative assessments have assessed numerous curriculum concepts covered throughout a unit of study. O'Connor (2002) emphasized that information from a varied summative assessment should be included when determining grades.

Due to the nature of the traditional grading practices implemented by both teacher A and B, formative assessments were not administered. In fact, both teachers relied solely on summative types of assessments to reflect the student classroom grade. If descriptive feedback had been given, the classroom grades represented in this study could have been higher, thus making the correlation analysis more significant.

O'Connor's Guideline # 5: Grade in pencil – keep records so they can be updated easily. Grades are not permanent until the student has demonstrated full understanding. In order to do this, students may need to take the summative assessment several times. Therefore, the teacher will need to offer students the option of retaking assessments. Wormeli (2006) pointed out that the retaking of assessments is conducted using consistent teacher guidelines. In addition, the teacher may reserve the right to offer a varied summative assessment as long as it is assessing the same curriculum concept. Guideline #5 further emphasizes allowing the students to retake summative assessments to provide more recent evidence of achievement. O'Connor (2002) believed

The most current information provides the most accurate depiction of student learning. If students demonstrate that past assessment information no longer accurately reflects their learning, that information must be dropped and replaced by the new information to accurately communicate student learning. (p. 128)

According to Guskey (1996), “Continuing to rely on past assessment data miscommunicates students’ learning” (p. 21).

The classroom grades represented in this study are the result of only one attempt to perform a task to accumulate enough points to earn a letter grade. Due to the traditional grading structure of the study high school, the grade a student received on a

task was the one that was recorded and averaged to eventually summarize the student's classroom grade. If students had been given a second opportunity to demonstrate their learning, the grade they earned in class could have been different than the grade represented in this study.

O'Connor's Guideline # 6: Crunch numbers carefully – if at all. The traditional practice is to use the mean of student performance to report student achievement. By using this practice, one low score can impact the final grade by misrepresenting actual student achievement. For example, if a student summative assessment score was 87, 89, 65, and 88, the mean would be 82. This numerical grade would not truly represent the level of achievement on three out of four summative assessments. O'Connor (2002) "emphasized final grades should (almost) never be determined by simply averaging the grades from several grading periods" (p. 144). Averaging scores falls far short of providing an accurate description of what students have learned. Guskey (1996) stated, "The purpose of grading and reporting is to provide an accurate description of what students have learned; the averaging must be considered inadequate and inappropriate" (p. 21). O'Connor suggested that teachers need to look at a "body of evidence" and use professional judgment in grading and reporting student achievement. This can be done by using the median or mode in determining final reporting. According to Wright (1994), "Grading by the median provides more opportunities for success by diminishing the impact of a few stumbles and by rewarding hard work" (p. 723).

The classroom grades used in this study were represented by using the averaging technique just described. This method could prove to not accurately communicate the learning that had taken place for the student. Because the classroom grade is a summary

of multiple concepts, the grade could have been different if a different technique had been used.

O'Connor's Guideline # 7: Use quality assessment(s) and properly recorded evidence of achievement. Using quality assessments to properly record student achievement is the principle behind Guideline #7. Traditionally in American education, for all core subject areas, educators have given assessments that have no clear target or purpose. In most cases, the assessments have been just recalling information from a story or event. The student (and in some cases the teacher) did not know the purpose for specific questions. In a standards based structure, every assessment has questions or performance rubrics that are aligned to specific curriculum standards. By having assessments aligned, teachers and students are able to more accurately determine knowledge or mastery of individual curriculum standards.

The assessments used to determine the grades used in this study were teacher made assessments. The researcher examined the classroom assessments and determined that questions on the assessments were not aligned to a specific concept or GLE. Once the test was graded, neither teacher A nor B knew what concepts had been mastered. The end result was the score being recorded in a grade book waiting to be averaged with the rest of the scores. If teachers knew the concepts that were missed and by whom, they could go back and reteach the students, giving them the opportunity to demonstrate their understanding. Demonstrating more understanding would have impacted their classroom grade.

O'Connor's Guideline # 8: Discuss and involve students in assessment, including grading, throughout the teaching/learning process. Involving students in the assessment

and grading process while teaching is beginning to occur. By being involved, students will have a better understanding of how their grades will be determined. In addition, they will have a better understanding of the expectations for reaching specific achievement levels and what they need to know or do to get there. Furthermore, if students are involved in the assessment and grading process, they will be better equipped to report the outcomes to parents and others.

There was no evidence that students were involved in the development of the grading process based on conversation between the two English III teachers and the researcher. Teachers in the study district taught the unit concepts and then announced the day the summative assessment would be administered. To some students, items appearing on the test may have come as a surprise if they were absent the day the concept was covered. This was determined by the researcher through observation and conversation that concepts were not identified at the beginning of the unit of study. Therefore, concepts were introduced on a random basis depending on the pacing of the curriculum.

Conclusion

This chapter reflected on the results uncovered by this study, which correlated the student achievement levels of the eleventh grade English III classroom grade to the achievement levels of the MAP scale score. These results were then related to the implementation of a new grading and reporting structure in a school. The leadership framework identified the findings of the McREL researchers Marzano et al. (2005). These researchers identified 21 leadership responsibilities that impact student achievement as it relates to ways a school leader could implement a new grading and reporting structure. Despite these responsibilities having varying degrees of significance,

they should be carried out by the school leader when implementing a change to a present structure such as changing the way a school grades and reports grading. The results section of this chapter focused on the eight guidelines used in developing a new grading and reporting system. These guidelines advocated by O'Connor (2002) were used to compare study results to changes that need to be made in developing a more effective grading and reporting system. To revise classroom grading and reporting, it is recommended that schools follow these guidelines. American educators have already begun to revise the curriculum, instruction, and assessment structures. Now it is time to examine traditional grading and reporting structures. By examining these existing structures, school districts will be able to better measure and communicate student achievement to not only the student and parents but to all stakeholders (e.g., community leaders, post-secondary educational institutions, government and military, and private corporations).

Current Status of Revising Grading and Reporting Structures in the Study District

At the completion of this study, this researcher, as the Assistant Superintendent of Curriculum and Instruction, recommended the Superintendent of the study district examine and revise the current grading and reporting structure. The decision was made by the study high school administrators to meet with a district committee comprised of teachers and administrators to act on this endeavor. The committee began by looking at current research surrounding accurate grading and reporting systems. Through this research, the committee examined the works of O'Connor, Marzano, Wormeli, and Guskey. This research led the committee to recommend changes to the current structure

beginning with the 2008-2009 school year. The following were the recommendations made by the committee:

1. Eighty percent of a student's grade should be summative in nature, while 20% of the classroom grade can consist of formative assessments such as homework and class work.
2. Students will be allowed to re-do two summative assessments per semester per subject area as long as the student abides by the district's re-do policy. The committee felt it was necessary to allow students this opportunity due to the classroom grade counting for eighty percent of the overall grade.
3. Teachers administering a summative assessment other than the district's approved summative assessment must have it approved by the Central Office administrator overseeing assessment. Questions on every proposed summative assessment must be aligned to either the Missouri Grade Level Expectations, Course Level Expectations, or the ACT Standards for Transition. In addition, each question must be assigned the appropriate Depth of Knowledge level. If a teacher proposes a summative performance project, the project must be accompanied by a scoring rubric with the same aforementioned alignment.
4. No extra credit of any type will be factored into the student's overall grade.
5. District report cards will remain unchanged for the 2008-2009 school year. The district report cards for grades K-5 already reflect standards-based grading measures. (study district, 2008, Spring)

Recommendations for Future Practice

Throughout the 2008-2009 school year, the grading and reporting committee continued to research and evaluate the grading structure. Based on a decision made in January 2008, the committee will make the following recommendations to be implemented for the 2009-2010 school year:

1. One-hundred percent of the student's overall grade will be summative.

Activities such as homework and class work will continue to be formative assessment but will not count toward the student's final grade. Formative assessments will be viewed as practice for the summative and will serve as report to guide and drive instructional planning for the classroom teacher.

2. Assigning students the grade of zero for not taking an assessment or completing a performance project will be further discussed. The discussion will determine if no zeros will be given or whether a grade of fifty percent will be assigned until a point in time during the grading period. After the time has expired, the student will then be assigned a failing grade for not completing the summative.

3. Report cards in grades K-5 will be revised to better reflect the standards taught in the curriculum. Report cards in grades 6-8 will remain the same but will be revised to reflect standards-based reporting for the 2010-2011 school year. In addition, during the revision process, non-academic reporting will be separated from the academic reporting. (study district, 2008, Spring)

As the study district continues its journey through the revision of its grading and reporting structure, further research will be conducted to determine how it can better

communicate more consistent results for its students. By better communicating more accurate levels of achievement, the study district students will be able to better identify their level of learning compared to where it needs to be for success in post-secondary education or the work place. Furthermore, officials from other school districts will be able to research and apply the study district model to their structure to produce more accurate and consistent grading and reporting for their students.

As school districts look to examine their grading and reporting systems, the researcher recommends that three steps need to occur. First, districts administrators should examine the current reality in which they are grading and reporting. Are they misrepresenting the students they are preparing to outside institutions? To determine if they are, a careful examination of data should be conducted. If the outcome of this examination reflects there is a need to change, then they can proceed on to the second step. The second step will need to take place after revising the grading and reporting system. Once the system has been changed, it needs to be monitored on a consistent basis by district administrators. Through this monitoring process, data will need to be analyzed between reporting with the old system and reporting with the new system. The outcome of this study will determine if student classroom achievement is being more accurately represented. Finally, in order to better assist school districts in determining successful prediction of student success, districts should develop a post high school tracking system. This system should involve a series of interest surveys before leaving high school and incremental years after they leave high school. It is recommended that school districts track their student's GPA for a period of five years after they graduate from high school.

Every time GPA data is collected, district administrators should compare the data to the GPA of the student upon graduating high school.

Implications of Research Findings

At the conclusion of this study, the researcher determined using the chi-square test that students' classroom grades did not indicate an accurate picture of their overall preparedness when compared to their achievement on a criterion-referenced test. All too often, students leave high school with high GPAs, feeling as if they are prepared to enter post-secondary educational institutions, when, in fact, they are not. The GPA is an averaged summary of all of the classroom grades a student has earned during high school. Because it is the final representation, it is meant to be an accurate reflection of the students' learning. It is also meant to be a predictor of the success the student is capable of achieving. If the current grading systems are operating on the assumptions that the GPA is an accurate predictor of student success, then schools are producing data that is misrepresenting students. Because of this misrepresentation, post-secondary institutions, governmental agencies, and perspective employers may believe they are hiring a person capable of performing at certain levels, when, in fact, they may not be capable.

The impact of this misrepresentation can be two fold. To begin with, if students have been led to believe they are capable of performing at a certain level and find out they are not, this could be a detriment to the students' self esteem. If this occurs, the outcome of students overcoming this obstacle will impact their ability to overcome obstacles in the future.

The second impact of student misrepresentation could be applied to students' place of employment. For example, if a business hires a student thinking they have hired

a person able to meet certain expectations, and the student is not able to meet those expectations, it could mean a loss of time and money. Furthermore, if students enroll in a post-secondary institution and are unable to perform as expected, they may not complete the course of study resulting in not only loss of money, but loss of confidence.

Recommendations for Future Research

The results of this study could be applied to other school districts without the limitations of this study (e.g., a school district with lower socioeconomic status). This study district had a free and reduced lunch student population of 25%. The same analysis could be applied to a study district with a free and reduced lunch student population of more than 50%. It is also recommended that this analysis be applied to areas other than Communication Arts. Assessment results from other areas such as Mathematics, Science and Social Studies could provide more substantive data. Furthermore, this analysis could be applied to a study district in the urban setting instead of a suburban setting. Finally, a recommendation for future research is to replicate this study in schools that have increased numbers of diversity in their student population.

References

- Airasian, P. W. (1994). *Classroom assessment* (2nd ed.). New York: McGraw-Hill.
- Bailey, J., & McTighe, J. (1996). Reporting achievement at the secondary level: What and how. In T. R. Guskey (Ed.), *Communicating student learning. Yearbook of the Association for Supervision and Curriculum Development* (pp. 119-140). Alexandria, VA: Association for Supervision for Supervision and Curriculum Development.
- Birk, L. (2000, January/February). Grade inflation: What's really behind all those A's? *Harvard Education Letter*, 1-5.
- Blount, H. P. (1997). The keepers of numbers: Teachers' perspectives on grades. *The Educational Forum*, 61(4) 329-334.
- Bracy, G. W. (2007, October). The first time everything changed. *Phi Delta Kappan*, 89(2), 119-136.
- Brookhart, S. M. (2004). *Grading*. Columbus, OH: Pearson Merrill Prentice Hall.
- Butler, R. (1988). Enhancing and undermining intrinsic motivation. *British Journal of Education Psychology*, 58, 1-14.
- Chappuis, S., Stiggins, R. J., Arter, J., & Chappuis, J. (2004). *Assessment for learning. An action guide for school leaders*. Portland, OR: Assessment Training Institute.
- Cross, L. H., & Frary, R. B. (1999). *Hodgepodge grading: Endorsed by students and teachers alike*. Paper presented at the annual meeting of the National Council on Measurement in Education, New York.
- CTB McGraw-Hill. (1999). *Missouri assessment program: Guide to test interpretation*. Monterrey, CA: Author.

- Danielson, C. (2002). *Enhancing student achievement: A framework for school improvement*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Dufour, R., & Eaker, R. (1998). *Professional learning communities at work* (2nd ed.). Bloomington, IN: National Educational Service.
- Dufour, R., Eaker, R., & DuFour, R. (2004). *Whatever it takes. How professional learning communities respond when kids don't learn*. Bloomington IN: Solution Tree.
- Durm, M. W. (1993, Spring). An A is not an A is not an A: A history of grading. *The Educational Forum*, 57, 294-297.
- Fraser, B. J., Walberg, H. J., Welch, W. W., & Hattie, J. A. (1987). Synthesis of educational productivity research. *Journal of Educational Research*, 11(2), 147-252.
- Garrison, C., & Ehringhaus, M. (n.d.). *Formative and summative assessment in the classroom*. Retrieved from the National Middle School Association Web site: http://www.nmsa.org/portals/0/pdf/publications/Web_Exclusive/Formative_Summative_Assessment.pdf
- Guskey, T. (1996). *Communicating student learning (the 1996 ASCD Yearbook)*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Guskey, T., & Bailey, J. M. (2001). *Developing grading and reporting systems for student learning*. Thousand Oaks, CA: Corwin Press Inc.
- Hunt, J. W. (2008). A nation at risk and no child left behind: Déjà vu for administrators. *Phi Delta Kappan*, 89(8), 580-585.

- Kohn, A. (2006). *The homework myth: Why our kids get too much of a bad thing*. Cambridge, MA: Da Capo Press.
- Littky, D. (2002). *The big picture. Education is everyone's business*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Maslow, A. H. (1943). A theory of human motivation. *Psychological Review*, 50, 370-396.
- Marzano, R. J. (2000). *Transforming classroom grading*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Marzano, R. J., Pickering, D. J., & Pollock, J. E. (2001). *Classroom instruction That works. Research based strategies for increasing student achievement*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Marzano, R. J. (2003). *What works in schools. Translating research into action*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Marzano, R. J. (2006). *Classroom assessment and grading that work*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Marzano, R. J., Waters, T., & NcNulty, B. A. (2005). *School leadership that works: From research to results*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Middleton, W. (1933). Some general trends in grading procedure. *Education*, 54(1), 5-10.
- Missouri Department of Elementary and Secondary Education. (2008). *Understanding your adequate yearly progress (AYP)*. Jefferson City, MO: Author.

- Missouri Department of Elementary and Secondary Education. (2004). *Missouri school improvement program: Standards and indicators manual*. Jefferson City, MO: Author.
- National Center for Educational Statistics. (2009). *NAEP - Frequently asked questions*. Retrieved February 16, 2009, from <http://www.nces.ed.gov/nationsreportcard/faq.asp>
- North Central Regional Educational Laboratory. (n.d.) *Summary of goals 2000: Educate America act*. Retrieved February 16, 2009, from <http://www.ncrel.org/sdrs/areas/issues/envrnmnt/stw/sw0goals.htm>
- O'Connor, K. (2002). *How to grade for learning: Linking grades to standards*. Arlington Heights, IL: Skylight Professional Development.
- Olson, L. (1995, June 14). Cards on the table. *Education Week*, 15(41), 23-28.
- Page, E. B. (1958). Teacher comments and student performance: A seventy-four classroom experiment in school motivation. *Journal of Educational Psychology*, 49(2), 173-181.
- Payne, R. K. (1996). *A framework for understanding poverty*. Highlands, TX: Aha! Processing.
- Popham, W. J. (2001). *The truth about testing. An educator's call to action*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Runyon, R. P., Coleman, K. A., & Pittenger, D. J. (2000). *Fundamentals of behavioral statistics*. Boston: McGraw-Hill Higher Education.
- Schmoker, M. (1999). *Results. The key to continuous school improvement* (2nd ed.). Alexandria, VA: Association for Supervision and Curriculum Development.

Stiggins, R. J. (1999). Assessment, student confidence, and school success. *Phi Delta Kappan*, 81(3), 191-198.

Stiggins, R. J., Arter, J., Chappuis, J., & Chappuis, S. (2006). *Classroom assessment for student learning. Doing it right – Using it well*. Portland, OR: Educational Testing Service.

United States Department of Education. (2003, May). No child left behind: Homework tips for parents. Retrieved February 21, 2009, from <http://www.ed.gov/parents/academic/involve/homework/homeworktips.pdf>

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

Wright, R. G. (1994). Success for all: The median is the key, *Phi Delta Kappan*, 14(5), 723-725.

Vitaé

Phillip David Kennedy was born in the District of Columbia on August 20, 1969, son of Gerald and Sara Kennedy. After graduating from Eureka High School in 1987, he completed his Bachelors of Science Degree in Secondary Education from Missouri State University in 1992. After serving four years as a secondary classroom teacher, he attended the University Missouri-St. Louis where he completed his Masters of Arts degree in Secondary Administration in 1998. While serving as a high school administrator, Phillip completed his Specialist Degree in Educational Leadership at Lindenwood University in 2006. Outside Phillip's devotion to improving student achievement in public education, he enjoys spending time with his family in the outdoors and participating in athletic activities. Currently, Phillip serves as the Assistant Superintendent of Curriculum and Instruction for the study district where he directs and monitors curriculum, instruction, assessment, and professional development.