

A COMPARISON OF EIGHTH GRADE READING SCORES BY STATE  
AND BY THE FOUR CENSUS-DEFINED REGIONS IDENTIFIED BY NAEP

by

WILLIAM R. GORDON II  
B.S. Rollins College, 1984  
M.S. Nova Southeastern University, 1989

A dissertation submitted in partial fulfillment of the requirements  
for the degree of Doctor of Education  
in the Department of Educational Research, Technology and Leadership  
in the College of Education  
at the University of Central Florida  
Orlando, Florida

Spring Term  
2009

Major Professor: Rosemarye Taylor

© 2009 William R. Gordon II

## ABSTRACT

This study provided information for policymakers and practitioners by comparing performance of eighth grade students in 2007 on state standardized reading assessments and by the four census-defined regions identified by NAEP. NCLB required states to set their own performance standards and to create their own data collection instruments resulting in increased transparency of student performance data and a lack of uniform accountability systems. The inability of educators, policy-makers, and the general public to make state-by-state comparisons in the area of reading was the catalyst for the study.

NAEP data were collected from NCES and state performance data were collected from the USDOE SY 2006-2007 CSPR to determine if a relationship existed between eighth grade students' state scores and NAEP scores in the four census-defined regions. Data were further disaggregated by low socioeconomic students and by nonwhite students.

A regression analysis was statistically significant in predicting: a) the state proficient and above scores from the NAEP proficient and above scores, b) the low socioeconomic state proficient and above scores from the NAEP proficient and above scores in the West census-defined region, and c) the nonwhite state proficient and above scores from the NAEP proficient and above scores in all regions. A regression analysis was not statistically significant in predicting low socioeconomic state proficient and above scores from the low socioeconomic NAEP proficient and above scores in the Midwest, South and Northeast regions.

In memoriam to  
Marcelyn Lorene Gordon  
The best Mom anybody could have asked for!

## ACKNOWLEDGMENTS

Several individuals have been especially helpful and supportive in the completion of this research. Dr. Rosemarye Taylor, Chair of my dissertation committee, was a constant source of encouragement and help as her genuine enthusiasm for the nature of my research was motivating and truly appreciated.

I was blessed to have an outstanding dissertation committee who provided accurate and timely insight and recommendations throughout this endeavor. Dr. George Pawlas, Dr. Walter Doherty and Dr. Lee Baldwin, are truly outstanding educators who demonstrate compassion and concern for their students for which I am truly grateful.

A title of one of my favorite songs by the Beatles “I Got By With A Little Help From My Friends” best explains how lucky I am to have such great support from so many people. Don Worcester and Julie Pepe helped enormously with the statistical computations. Patty Glore, Judy Savage, Jill Dunn, Susan Gluckman and Dr. Michele Green were willing to serve as a “second pair of eyes” and to help edit countless pages of writings and research. My Assistant Principal team: Kimrey Sheehan, Dave Stanley, Michael Byrne, DeMikki Joiner, and Heather Hilton held down the fort and propped me up throughout this process. They are true colleagues! Additionally, I am most appreciative of having a supportive boss, Dr. Cathy Pope, whose encouragement and support was unparalleled.

Lastly, my family, were the utmost and staunchest supporters of all. My father, William Gordon, the Vice President Emeritus of Rollins College, has served as a mentor

and coach for me throughout my entire life. He is truly my best friend! My mother, Peggy Gordon, who is deceased, would be so proud of her son and I dedicate my efforts and this book in her name. My aunt, Mary Clerk, the successful business owner, acted as a second mom and kept me going. My children, Trey and Whitney, are the best and served as an inspiration for me to complete this project. And finally, my wife, Patty, was the greatest supporter of all. Her unselfish and unyielding support is the main reason I was able to earn my degree. I love her very much and thank her tremendously for all she does for our family.

## TABLE OF CONTENTS

LIST OF FIGURES .....	xii
LIST OF TABLES .....	xiii
CHAPTER ONE: INTRODUCTION.....	1
Review of Literature .....	3
Sputnik Crisis.....	4
Elementary and Secondary Act of 1965 .....	5
A Nation at Risk.....	6
Standardized Testing.....	8
The No Child Left Behind Act of 2001(NCLB).....	11
Purpose of the Study .....	14
Research Questions.....	14
Definition of Terms.....	15
Study Design.....	17
Population and Sample .....	17
Instrumentation .....	17
Reliability.....	18
Data Collection .....	18
Data Analysis .....	19
Assumptions.....	20
Limitations .....	20
Significance of the Study .....	20

CHAPTER TWO: REVIEW OF LITERATURE.....	23
Introduction.....	23
Sputnik Crisis.....	24
Elementary and Secondary Act of 1965 .....	25
A Nation at Risk.....	26
Improving America’s Schools Act and Goals 2000: Educate America Act.....	28
Standardized Testing.....	31
The National Assessment of Educational Progress .....	35
The No Child Left Behind Act of 2001(NCLB).....	39
Title 1: Improving the Academic Achievement of the Disadvantaged.....	41
Title VI: Flexibility and Accountability .....	47
Summary.....	50
CHAPTER THREE: METHODOLOGY .....	54
Introduction.....	54
Statement of the Problem.....	55
Research Questions.....	55
Population and Sample .....	56
Instrumentation .....	56
Instrument Reliability and Validity .....	56
Background.....	56
Reliability.....	57
Data Collection .....	58
Data Analysis .....	58



Summary .....	59
<b>CHAPTER FOUR: ANALYSIS OF DATA .....</b>	<b>60</b>
Introduction.....	60
Purpose of the Study .....	62
Research Questions.....	62
Demographics .....	63
Analysis of Data.....	64
Research Question 1 .....	64
Research Question 2 .....	66
Research Question 3 .....	69
Low Socioeconomic Census-Defined Region (Midwest).....	69
Low Socioeconomic Census-Defined Region (Northeast).....	71
Low Socioeconomic Census-Defined Region (South) .....	73
Low Socioeconomic Census-Defined Region (West) .....	75
Research Question 4 .....	78
Nonwhite Census-Defined Region (Midwest).....	78
Nonwhite Census-Defined Region (Northeast) .....	81
Nonwhite Census-Defined Region (South) .....	84
Nonwhite Census-Defined Region (West).....	87
Summary .....	89
<b>CHAPTER FIVE: CONCLUSIONS AND RECOMMENDATIONS .....</b>	<b>90</b>
Introduction.....	90

Statement of the Problem.....	92
Research Questions.....	92
Data Collection .....	93
Instrument Reliability and Validity .....	93
Background.....	93
Reliability.....	94
Summary and Discussion.....	95
Research Question 1 .....	95
Discussion.....	96
Research Question 2 .....	97
Discussion.....	97
Research Question 3 .....	99
Low Socioeconomic Census-Defined Region (Midwest).....	99
Low Socioeconomic Census-Defined Region (Northeast).....	100
Low Socioeconomic Census-Defined Region (South) .....	100
Low Socioeconomic Census-Defined Region (West) .....	101
Discussion.....	102
Research Question 4 .....	103
Nonwhite Census-Defined Region (Midwest).....	104
Nonwhite Census-Defined Region (Northeast) .....	105
Nonwhite Census-Defined Region (South) .....	106
Nonwhite Census-Defined Region (West).....	107
Discussion.....	108
Conclusions.....	109

Implications for Practice .....	112
Future Research .....	113
APPENDIX A: NCES 2007 STATE SNAPSHOT REPORT (ALABAMA) .....	114
APPENDIX B: SY 2006-2007 CSPR REPORT (ALABAMA) .....	116
APPENDIX C: THE FOUR CENSUS-DEFINED REGIONS OF NAEP .....	118
APPENDIX D: SPSS WORKSHEET .....	120
APPENDIX E: STATE AND NAEP DATA (PROFICIENT AND ABOVE) .....	122
APPENDIX F: STATE AND NAEP BY CENSUS-DEFINED REGION .....	124
APPENDIX G: LOW-SOCIOECONOMIC CENSUS-DEFINED REGION .....	126
APPENDIX H: NONWHITE CENSUS-DEFINED REGION .....	128
APPENDIX I: HISTOGRAMS WITH NORMAL CURVES (MIDWEST REGION) .....	130
APPENDIX J: HISTOGRAMS WITH NORMAL CURVES (N.E. REGION) .....	132
APPENDIX K: HISTOGRAMS WITH NORMAL CURVES (SOUTH REGION) ....	133
APPENDIX L: HISTOGRAMS WITH NORMAL CURVES (WEST REGION) .....	135
APPENDIX M: STUDENT PERFORMANCE ON STATE AND NAEP EXAMS ....	137
APPENDIX N: UCF SUBMISSION OUTCOME LETTER .....	139
LIST OF REFERENCES .....	141

## LIST OF FIGURES

Figure 1. The Four Census-Defined Regions of NAEP.....	19
Figure 2. Scatter Plot of State Proficient and Above and NAEP Proficient and Above...	65
Figure 3. Scatter Plot of Midwest Census-Defined Region (Low Socioeconomic) .....	70
Figure 4. Scatter Plot of Northeast Census-Defined Region (Low Socioeconomic).....	72
Figure 5. Scatter Plot of South Census-Defined Region (Low Socioeconomic).....	74
Figure 6. Scatter Plot of West Census-Defined Region (Low Socioeconomic).....	76
Figure 7. Scatter Plot of Midwest Census-Defined Region (Nonwhite) .....	79
Figure 8. Scatter Plot of Northeast Census-Defined Region (Nonwhite).....	82
Figure 9. Scatter Plot of South Census-Defined Region (Nonwhite).....	85
Figure 10. Scatter Plot of West Census-Defined Region (Nonwhite) .....	87

## LIST OF TABLES

Table 1 Model Summary of State and NAEP (Proficient and Above) .....	65
Table 2 Regression Model of State and NAEP (Proficient and Above).....	66
Table 3 Wilcoxon (Midwest).....	67
Table 4 Wilcoxon (Northeast) .....	68
Table 5 Wilcoxon (South).....	68
Table 6 Wilcoxon (West).....	68
Table 7 Model Summary of Low Socioeconomic (Midwest) .....	71
Table 8 Model Summary of Low Socioeconomic (Northeast).....	73
Table 9 Model Summary of Low Socioeconomic (South) .....	75
Table 10 Model Summary of Low Socioeconomic (West) .....	77
Table 11 Regression Model of Low Socioeconomic (West).....	78
Table 12 Model Summary of Nonwhite (Midwest).....	80
Table 13 Regression Model of Nonwhite (Midwest) .....	81
Table 14 Model Summary of Nonwhite (Northeast) .....	83
Table 15 Regression Model of Nonwhite (Midwest) .....	84
Table 16 Model Summary of Nonwhite (South) .....	85
Table 17 Regression Model (South) .....	86
Table 18 Model Summary of Nonwhite (West) .....	88
Table 19 Regression Model of Nonwhite (West) .....	89

## CHAPTER ONE: INTRODUCTION

With the advent of the No Child Left Behind Act of 2001 (NCLB), the idea of a national assessment system sparked debate in education circles. Due to an increased transparency of student performance data and the lack of uniform accountability systems, the inability of educators, policy-makers and the general public to make state-by-state comparisons in the area of reading was apparent (National Center for Educational Statistics, 2007). As this all-encompassing education policy continues to influence and reform public education, a standardized accountability system may provide a legitimate way for state-by-state comparisons to be made.

The implementation of mandates at the school level, in relation to the No Child Left Behind Act of 2001 (NCLB), presents educators with tremendous challenges. The premise of the NCLB was set on four pillars: accountability for results, emphasis on doing what works based on research, expanded parental options, and expanded local control and flexibility (United States Department of Education, 2004a). However, the main focus of the NCLB was on accountability. “The No Child Left Behind Act of 2001 (Public Law 107-110) sets demanding accountability standards for schools, school districts, and states, including new state testing requirements designed to improve education” (Lunenburg & Ornstein, 2004, p. 453). Other goals of the NCLB include the following: all Limited English Proficient students will be proficient in English, reading and math, by the 2005-06 school year, all teachers were to be highly qualified; all students were to attend school in a safe and drug-free learning environment; all students

were to graduate from high school; and by the 2013-14 school year, all students were to be proficient in reading and math (Thompson, 2004).

The NCLB requires states to set their own performance standards and to create their own data collection instruments. The inability of educators, policy-makers and the general public to make state-by-state comparisons in the area of reading was apparent (National Center for Educational Statistics, 2007). The National Assessment of Educational Progress (NAEP) can play an important role in confirming statistics and in allowing comparisons to be made state by state. The Ad Hoc Committee on Confirming Test Results (2002) reports that “differences between the National Assessment of Educational Progress and state testing programs must be explored and reported” (Stoneberg, 2007). In addition, Stoneberg suggests comparisons can be made between NAEP’s “At or Above Basic” data and a state’s Adequate Yearly Progress (AYP) results. The use of trend lines produced from NAEP data and state-created exams can provide a reliable method of confirming state AYP results (Stoneberg, 2007).

When the 2005 NAEP results in eighth grade reading were reported by the National Center for Education Statistics (NCES), a large gap in achievement between the two types of assessments became evident (National Center for Educational Statistics, 2007). This disparity was attributed to the stringency of standards by individual states as well as to the content of the standards they chose to measure. Currently, accurate comparisons cannot be made between state proficiency standards and NCLB standards because states were free to design assessments and to determine standards they wish to measure (United States Department of Education, 2004a). If a national assessment system

was implemented in all states, one that can determine if the goals of the NCLB were being met, comparisons were made.

### Review of Literature

In 1858, our nation's Department of Education in its "Twentieth Report" declared, "Our system of free schools was sustained directly by the people, without special care or direct aide from the government" (Twentieth Report found in Katz, 1968, p. 3). At that time, society's sentiment toward schools was favorable, and the responsibility for augmentation of any deficiencies which may have existed was placed with the family of each individual student (Katz, 1968).

As our country evolved from a primarily agrarian society into an urban industrialized nation, new challenges in educating the masses appeared. Due to an increasing population and a disbanding homogeneous society, our public education system began to lose efficiency and effectiveness. During the first half of the 20<sup>th</sup> century, our nation's Department of Education responded by implementing structural reform that resulted in a "system" which has helped to secure the right for the masses to have a free education (Katz, 1968).

In recent decades the call to reform public education has been the mantra of politicians, private-sector entrepreneurs, parents, and a multitude of others. Understandably, the wide-range of expectations of stakeholders and the fact that schools have become multipurpose institutions expected to address every social, economic and political ill ensure that some aspect of society will always be unhappy with public



education (Reese, 2007). Even so, schools remained as critical entities and were indispensable to a republic (Tyack, 2000).

### Sputnik Crisis

The launch of Sputnik by the Soviet Union in the late 1950s, raised concern with the American public that the Soviet school system was superior and would produce superior scientists. Consequently, American public schools received seemingly endless negative press regarding their ability to compete on an international level (Reese, 2007). Press reports indicated public schools were not preparing children in the areas of mathematics and science and insinuated that our nation's education system was in need of a serious overhaul. The United States Department of Education's "life adjustment education" curriculum, which sought to make school relevant to teenagers, was picked apart. Speeches and interviews given by politicians, military leaders and scientists called for a return to a more traditional form of education. Sputnik set a nasty precedent in which public education received all the blame for America's social crisis; yet, when the crisis was resolved, it received no credit (Bracey, 2007).

In 1959, Vice Admiral Rickover, in his book, *Education and Freedom*, declared: "Our schools are the greatest cultural lag we have today" (Rickover, 1959, p. 23). As a response to the mounting crisis, Congress poured substantial amounts of money into public education. Accordingly, President Eisenhower signed into law the National Defense Education Act, which bolstered science and math offerings in the school setting (Mondale & Patton, 2001).

### Elementary and Secondary Act of 1965

In the early 1960's, dissatisfaction with American schools continued and for the first time, a school accountability movement, due to the large amount of public funds being allocated, was propagated (Bowers, 1991). Public perception of schools had never fully recovered from the Sputnik Crisis, and newly designed large standardized tests such as the National Assessment of Educational Progress (NAEP) showed students still lacked basic skills (Vlaanderen, R. as cited in Bowers). Additionally in the 1960s, the Civil Rights Movement had been roiling under the surface due to racial inequalities, including those in the area of education.

In 1964, Congress enacted the Civil Rights Act, which banned racial discrimination in all federally funded institutions and gave the United States Department of Education the authority to collect racial data from schools (Brown, 2004). Pressure mounted to increase educational opportunities for all children, and the Elementary and Secondary Act of 1965 (ESEA) was established. At the time, ESEA was considered the most monumental effort the federal government had attempted in its efforts to influence state and local boards to reform their practices (Kantor, 1991). Significant monies were included to entice states and school districts to align to ESEA policies. Over four billion dollars in aid (ie, Title I funds) was offered to assist disadvantaged youth in the educational process (Mondale & Patton, 2001). Subsequent to assisting disadvantaged youth, funding was also set aside to assist in the areas of education technology, professional development, class-size reduction, safe and drug free schools, bilingual

education, Native American education, charter schools and head start (Electronic Summary of The Elementary and Secondary Education Act of 1965, 2008).

A major expectation of the ESEA was that schools using governmental funds were expected to improve academic performance of at-risk students in the areas of reading and math, when compared to other students who were not at risk, throughout their state (Wong & Meyer, 1998). In addition, the ESEA was to be evaluated every five years to insure goals were met (Electronic Summary of The Elementary and Secondary Act of 1965, 2008).

### A Nation at Risk

In 1969, the Gallup organization began to influence national opinion of public education by producing surveys regarding education. These surveys, while largely reflections of public opinion, brought further criticism to school systems. Gallop survey data showed that in 1978, 41 percent of respondents felt public schools were not as good as they used to be, while only 35 percent felt they had improved (Tyack & Cuban, 1995).

With national opinion of public education again suggesting students were not as well educated as their parents; an effort to reform education was instituted by Congress. In 1981, Secretary of Education T. H. Bell was directed by Ronald Reagan to create the National Commission on Excellence in Education, which linked our country's poor economic conditions to a failing public education system (Tyack & Cuban, 1995, p. 14). The commission was directed to pay special attention to teenage youth while critiquing our public schools and universities (Electronic Summary of *A Nation at Risk*, 2001).

In addition to the commissions being accountable for teenage youth, it was directed to report on the following: the status of teaching and learning in our nation's public and private schools, colleges, and universities; the performance of American schools and colleges in relationship to other advanced nations; studying student achievement in high school as it related to college admissions; identifying educational programs, which result in notable student success in college; assessing how major social and educational changes in the last quarter century impacted student achievement; and identify and define<sup>3</sup> future considerations that need to be overcome to achieve excellence in education (Electronic Summary of a *Nation at Risk*, 2001).

*A Nation at Risk* brought to light the considerable decline of our nation's school system. In the report, the commission reiterated that the positive gains made from reform achieved after the Sputnik crisis had all but been diminished. Furthermore, the commission stated that the multiple demands which other social, personal and political institutions did not tackle or solve should not continue to be placed on the nation's schools and universities. Ultimately, these demands detracted from the purpose of the educational system (Electronic Summary of *A Nation at Risk*, 2001).

Because of America's declining economic status and the thriving economies of our old World War II foes, angst about the perceived decline in public education resulted in President Ronald Reagan calling for reform. Reagan suggested that not only was the problem economic but that the enforcement of civil rights legislation had contributed to the decline in basic education. Many educators reputed these claims (Mondale & Patton, 2001). Author, Nicholas Lemann supported their view (as cited in Mondale & Patton,

2001), stating that NAEP scores showed the decline was not as dramatic as reported in *A Nation at Risk*. Although public education had many proponents, the quality of America's public schools continued to be questioned by the public.

### Standardized Testing

Assessing student performance and the performance of teachers has been a driving force behind educational reform movements since the 1970s. Policymakers have learned that test results can be used as a vehicle to implement educational reform (Madaus, 1985). Consequently, standardized tests, due to their relatively inexpensive implementation, their ability to be easily mandated, their flexibility and ease of implementation, and their transparency, provided a mechanism through which reformers rallied their support in an effort to overhaul our nation's public education system (Linn, 2000).

Standardized tests came into existence just before the turn of the twentieth century. Initially, standardized tests were not used at the national level to compare student performance. Instead, they were used as means to compare teacher performance within schools, between schools, and between districts and to see how students performed on outside exams (Resnick, 1980).

By World War I, test designers began to provide data that enabled educators to make comparisons on a more global scale. They developed exams that presented data about representative populations and allowed students to be placed into homogeneous groups. Although data was not always used effectively, the exams did ensure they were

implemented with some form of uniformity by having directions for conformity of implementation and interpretation and by norming grading criteria (Resnick, 1980).

The role of standardized testing was diminished from 1930 through the 1950s. During this era, Americans had confidence in the public education system (Resnick, 1980). In the 1960s, politicians and reformists called for all children to have equal access and equal educational opportunities in our nation's public education system. A joint committee of legislators was formed to promote these ideals, and the Elementary and Secondary Education Act of 1965 (ESEA) was established to provide financial assistance to schools and districts with substantial numbers of low-income students. In addition, the federal government contracted with the Education Commission on States and created the National Assessment of Educational Progress (NAEP) in 1969. The goal of the establishment of NAEP was to monitor achievement in 10 learning areas and to assess change in achievement over the years. Information gathered from NAEP was to then be reviewed and used by policymakers to institute change in public education (Resnick, 1980).

ESEA funded a variety of programs to facilitate widespread reform in the public education system. ESEA consisted of six titles. The most significantly funded of these was Title I, which held a provision for required testing of student performance as a means of evaluating funded programs (Resnick, 1980). The significant funding led to a boon in the testing industry, and norm-referenced testing with Normal Curve Equivalent (NCE) was introduced. Essentially, NCEs were standardized with a mean score of 50 and a standard deviation of 21.06 (Linn, 2000).

To accomplish the implementation of NCEs, the Title I Evaluation and Reporting System (TIERS) was established, and testing companies experienced large-scale financial success (Tallmadge & Wood, 1981). TIERS data were collected and information presented to the general public painted a relatively positive picture. Elizabeth Reisner and her colleagues, who were commissioned to study TIERS in 1982, reported: 1) the program provided the achievement data requested by Congress as specified, 2) the technical assistance and evaluation model had a positive effect on instructional practices at the local level, and 3) new approaches to the evaluation process were introduced to educators (Reisner, et al, 1982).

While TIERS appeared to be successful in evaluating and reporting NCEs, other studies produced different opinions about the success of TIERS. Linn and his colleagues found a number of variables such as student selection, scale conversion errors, administration conditions and teaching to the test to be among factors which contributed to potential corruption of the data and inflated results (Linn, et al. 2002)

Minimum Competency Exams (MCTs), as a means for large scale standardized testing, became the norm in the 1970s and 1980s. States, such as Florida, were closely studied due to their requiring students to pass their version of an MCT, the High School Competency Test (HSCT), in order to graduate (Linn, 2000). According to Linn (2000), at the inception of the HSCT in 1977, white and Hispanic students performed well on their first attempt to pass. African-American students performed considerably lower, but their scores did increase as the test was implemented over time (Linn, 2000). The underlying lesson learned from the long-term implementation of the HSCT and other

MCTs was that student performance becomes flat over time. Straight-lined standardized test data over long periods of time should be reviewed with caution when using it to drive policy (Linn, 2000).

In the 1990s, the promotion of higher standards by federal and state legislatures resulted in a standards-based reform movement and the development of standards-based assessments. Swanson and Stevenson (2002), in their article, “Standards-Based Reform in Practice: Evidence on State Policy and Classroom Instruction from the NAEP State Assessment” describe school reform efforts in the 1970s and 1980s as largely unsuccessful due to their lack of connection between national policy and the classroom. On the other hand, they state that standards-based reform, when effective standards-based assessments were included, can lead to effect improvement of instructional practices (Swanson & Stevenson).

According to Blum (2000), instituting higher content standards and establishing a standards-based assessment system allowed states to set benchmarks and performance standards for student achievement. By the mid-to-late 1990s, almost every state in the nation had adopted some form of standards-based assessment (Blum).

#### The No Child Left Behind Act of 2001(NCLB)

On January 8, 2002, President George W. Bush signed Public Law 107-110, The No Child Left Behind Act of 2001 (NCLB), into law. This largely bipartisan effort of Congress was essentially the reauthorization and expansion of the Elementary and Secondary Act of 1965. NCLB represents Congress’ effort to bring greater accountability



and improvements in student achievement to the forefront in the educational reform movement (Shaul & Ganson, 2005).

The premise of NCLB was set on four pillars: accountability for results, emphasis on doing what works based on scientific research, expanded parental options, and expanded local control and flexibility (United States Department of Education, 2004a). Under NCLB, states were given more local control and were expected to close the achievement gap for all children, including those from economically disadvantaged families. Linn noted, NCLB's emphasis on promoting learning in groups of students who have lagged behind was positive. By focusing on standards based reform and promoting higher state standards, NCLB increased opportunities for students. Conversely, NCLB was driven by accountability and relied too heavily on high-stakes tests that focus primarily on reading and mathematics. In fact, NCLB did not allow states to use other indicators that were not outlined in Sec. 1111 (b)(2)(A)(i) because they could outweigh how students perform in these areas (Linn, 2003).

Another major provision, Sec. 1111 (b)(1)(A) of NCLB, was the expectation that states set demanding content standards as well as challenging student achievement standards (Linn, 2003). States were required to establish measurement tools to evaluate whether Local Educational Agencies (LEAs) were making AYP within parameters set in Title 1 (United States Department of Education, 2004d). Determining AYP was the means by which a state measured its demanding content and challenging student achievement standards. Each state established its own criteria when assessing a student's yearly progress and must report these scores to parents. The federal government required

states to have 100% of their students demonstrate success on state developed standardized tests by the year 2014 (United States Department of Education, 2004d).

States that receive federal funds under NCLB were required to have fourth and eighth grade students participate in NAEP reading and mathematics assessments every two years. According to the U.S. Department of Education, “the NAEP provides parents with a wealth of data about the condition of education in the United States” (United States Department of Education, 2004d).

The policies of NCLB were stated in the following ten titles of NCLB: Title I – Improving the Academic Achievement of the Disadvantaged; Title II – Preparing, Training, and Recruiting High Quality Teachers and Principals; Title III – Language Instruction for LEP and Immigrant Students; Title IV – 21<sup>st</sup> Century Schools; Title V – Promoting Informed Parental Choice; Title VI - Flexibility and Accountability; Title VII - Indian, Native Hawaiian, and Alaska Native Education; Title VIII - Impact Aid Program; Title IX - General Provisions; Title X -Repeals, Re-designations, and Amendments to Other Statutes (107<sup>th</sup> Congress, 2002).

In creating NCLB, the federal government’s intent was to influence a change in the way our state public school systems conduct business. Although all of the ten titles were important, for the purpose of the study, the researcher has chosen two titles, Title I and Title VI, due to their pertinence.

## Purpose of the Study

The purpose of this study was to provide comparisons of the proficiency of eighth grade students in reading on the NAEP and state standardized tests in 2007. The researcher chose to compare all states due to an NCLB requirement that all states report eighth grade student reading proficiency data and the availability of NAEP scores of eighth grade students by state. The desired outcome of this study was to provide important information so that more accurate comparisons of students' performance can be made on a national level in order to validate the provisions of NCLB.

## Research Questions

The study was guided by the following research questions:

1. What relationship exists, if any, between the percent of eighth grade students identified as proficient in reading on state assessments and on the NAEP in 2007?
2. What comparison can be made, if any, between the percent of eighth grade students identified as proficient and above in reading on state assessments and on the NAEP, in the four census-defined regions identified by the NAEP in 2007?
3. What relationship exists, if any, between the percent of low socioeconomic eighth grade students identified as proficient and above in reading on state assessments and on the NAEP, in the four census-defined regions identified by the NAEP in 2007?
4. What relationship exists, if any, between the percent of nonwhite eighth grade students identified as proficient and above in reading on state assessments and on the NAEP, in the four census-defined regions identified by NAEP in 2007?

## Definition of Terms

For the purpose of this study, the following definitions were utilized to clarify terminology:

Adequate Yearly Progress (AYP) -- AYP was a provision in NCLB that requires schools to test 95% of their students. In addition, subgroups were broken down into race, students with disabilities, economically disadvantaged students, and students with limited English proficiency (LEP) groups.

All Students -- Students or children from a broad range of backgrounds and circumstances, including disadvantaged students and children, students or children with diverse racial, ethnic, and cultural backgrounds, American Indians, Alaska Natives, Native Hawaiians, students or children with disabilities, students or children with limited-English proficiency, school-aged students or children who have dropped out of school, migratory students or children, and academically talented students and children.

The Elementary and Secondary Act of 1965 (ESEA) -- ESEA was the most expansive federal government bill ever approved by Congress to improve education. Its main focus was to address the issue of inequality in education. Established in 1965, it was reauthorized regularly.

Limited English Proficiency (LEP) -- Students who have not acquired English as their first language.

Low Socioeconomic Students -- Students from low income families that do not have enough money to access materials necessary to promote and support learning.

Minimum Competency Exams (MCES) -- MCES were established during the 1970s as a means to determine if students had met a level of competence in mathematics, reading and writing.

National Assessment of Educational Progress (NAEP) -- NAEP was the Nation's Report Card that has provided statistical data since 1969 in the area of reading and mathematics.

National Center for Education Statistics (NCES) -- NCES collects, analyzes and makes available data related to education in the US and other nations.

Nonwhite Students – Students who were tested on the NAEP and state standardized tests that are not white.

The No Child Left Behind Act of 2001 (NCLB) -- The NCLB was essentially the reauthorization of the Elementary and Secondary Act of 1965. The policies of the law were outlined throughout the following ten titles of NCLB: Title I – Improving the Academic Achievement of the Disadvantaged; Title II – Preparing, Training, and Recruiting High Quality Teachers and Principals; Title III – Language Instruction for Limited English Proficient (LEP) and Immigrant Students; Title IV – 21<sup>st</sup> Century Schools; Title V – Promoting Informed Parental Choice; Title VI - Flexibility and Accountability; Title VII - Indian, Native Hawaiian, and Alaska Native Education; Title VIII - Impact Aid Program; Title IX - General Provisions; Title X - Repeals, Re-designations, and Amendments to Other Statutes.

Title -- The areas under NCLB which house the actual policies and procedures developed by the federal government.

Title 1 Evaluation and Reporting System (TIERS) -- Still in existence, TIERS was established under ESEA to collect achievement data for congress and to provide technical assistance at the local level.

## Study Design

### Population and Sample

Eighth grade student reading performance data on state assessments and the NAEP in 2007 were targeted for this study. As a part of the requirements of NCLB, states were required to report eighth grade reading performance data and were required to have students participate in the NAEP. NAEP scores were reported by the United States Department of Education. State scores were available for constituents of the state and were reported to the United States Department of Education.

### Instrumentation

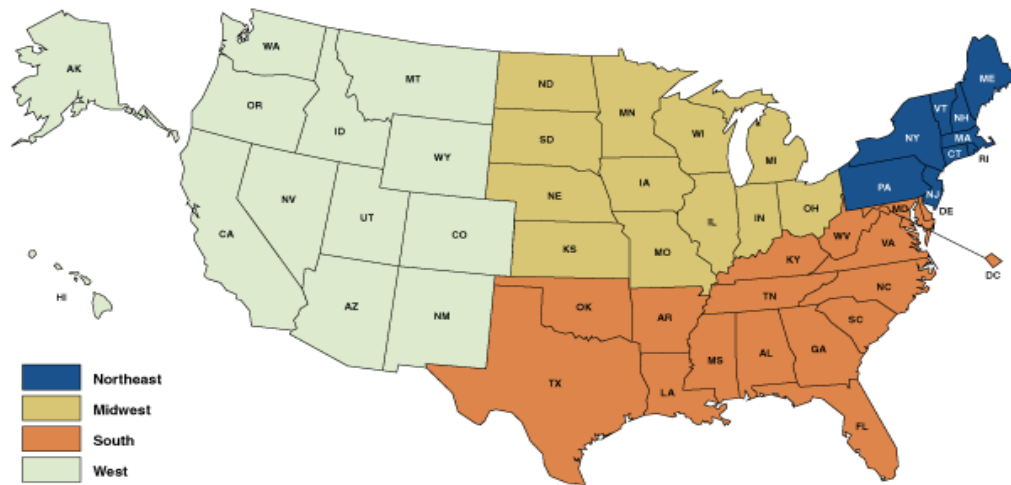
The researcher collected NAEP data from the National Center for Education Statistics, a well respected division of the United States Department of Education's Institute of Education Services. State performance data were collected from the United States Department of Education SY 2006-2007 Consolidated State Performance Report. Collected data were analyzed through a series of statistical procedures to determine if a relationship existed and comparisons could be made between the NAEP and state standardized reading scores in 2007

### Reliability

According to Jones and Olkin (2004), NAEP was considered the “Gold Standard” in measuring student achievement. Education Testing Service (ETS) was responsible for the design, analysis and reporting of the 2007 NAEP reading data. NAEP assessments contained both multiple-choice and constructed response questions. Constructed response questions required students to provide their own answers. In order to score large numbers of constructed response questions, over 3 million annually, with a high level of accuracy, NAEP incorporated extensive quality control measures to ensure reliability (United States Department of Education, 2008a). It was incumbent on each state to develop its own procedures to assure accurate and reliable design, analysis and reporting of state assessments under the terms of NCLB.

### Data Collection

In November 2008, eighth grade reading performance data were collected from the NCES 2007 State Snapshot Report (Appendix A) and from SY 2006-2007 Consolidated State Performance Reports (Appendix B). Collected data were organized in an SPSS worksheet. Using SPSS, data were disaggregated by the four census-defined regions identified by NAEP as shown in Figure 1. Upon completion of regional disaggregation, data were further disaggregated by low socioeconomic and by nonwhite students. This information was analyzed using several statistical procedures to answer the four research questions.



- *Northeast:* Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont;
- *South:* Alabama, Arkansas, Delaware, District of Columbia, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia, and West Virginia;
- *Midwest:* Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, and Wisconsin; and
- *West:* Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington, and Wyoming.

Source: U.S. Department of Education, Institute of Education Sciences, National Center For Education Statistics, National Assessment of Educational Progress (NAEP), Census-Defined Regions of NAEP

Figure 1. The Four Census-Defined Regions of NAEP

### Data Analysis

The following 2007 demographic data were collected within an SPSS worksheet:

- (a) name of state, (b) the four census-defined regions of NAEP, (c) the percent of eighth grade students identified as proficient and above in reading on NAEP, (d) the percent of low socioeconomic eighth grade students identified as proficient and above in reading on state assessments, (e) the percent of low socioeconomic eighth grade students identified



as proficient and above on NAEP, (f) the percent of nonwhite eighth grade students identified as proficient and above in reading on state assessments, and (g) the percent of nonwhite eighth grade students identified as proficient and above in reading on NAEP.

### Assumptions

The first assumption was the data collected from the National Center of Educational Statistics on NAEP eighth grade reading scores were accurate and reliable. The second assumption was the data collected from the United States Department of Education SY2006-2007 Consolidate State Performance Report were accurate and reliable.

### Limitations

The following study will be limited as follows:

1. The information provided for the study was provided by another agency.
2. State standards and assessments were designed by each state and had varying degrees of difficulty and that proficiency was set at varying degrees of difficulty.

### Significance of the Study

NCLB was essentially the reauthorization of the Elementary and Secondary Act of 1965. The policies of the law were outlined throughout the ten titles of NCLB. Although all ten titles were important, two have significant impact: Title I - Improving

the Academic Achievement of the Disadvantaged and Title VI - Flexibility and Accountability.

Under NCLB, the accountability provisions of Title I were strengthened in 2001 to provide a mechanism for holding states, school districts, and schools accountable for improving the academic achievement of all students and turning around low-performing schools. Additionally, it required that students in low-performing schools were to be afforded alternatives to enable those students to receive a high-quality education. Schools and districts had to demonstrate all students were meeting annual targets established by states through their assessment systems. Annual assessments were required at grades 3-8 and once in grades 10-12. Adequate Yearly Progress (AYP) provisions were housed under Title I. In order for schools to obtain the AYP provision, 95% of students in the school must be tested. In addition, states were required to report subgroup performance by race, students with disabilities, economically disadvantaged students, and students with Limited English Proficiency (LEP). All students were expected to be proficient in reading and mathematics by the 2013- 2014 school year.

The principal goal of creating Title VI, by the federal government, was to create collaborative environments in which states, Local Education Agencies (LEAs) and higher education institutions develop assessments in order to elevate student performance and close the achievement gap. Section 6111 of Sub Part 1 of the Accountability section of Title VI houses the majority of the provisions regarding the federal government's desire to influence states to align their accountability systems to academic content and standards. By passing this legislation the federal government provided funds, in the form

of grants, to promote the creation of high academic standards and valid assessment by states.

NCLB was legislated at the federal level and states were given the autonomy to develop their own curriculum standards and design their own assessments to measure student performance on the standards. If the state assessments measured different standards, accurate comparisons can not be made.

Few studies in the United States compared eighth grade reading performance of students on state level exams and the NAEP. This study sought to determine what relationship, if any, existed between eighth grade reading performance on state assessments and the NAEP. Additionally, data were disaggregated by the four census-defined regions of NAEP. Data were further disaggregated by low socioeconomic and by nonwhite students.

It was an assumption of the researcher that different states' exams had varying degrees of difficulty and that proficiency was set at varying degrees of difficulty which, in turn, did not allow accurate comparisons to be made between states. Results of this study could influence national policy and promote the need for national standards and a national assessment system as a means to mitigate the disparities presented in state reports.

## CHAPTER TWO: REVIEW OF LITERATURE

### Introduction

In 1858, our nation's Department of Education in its "Twentieth Report" declared, "Our system of free schools was sustained directly by the people, without special care or direct aide from the government" (Twentieth Report found in Katz, 1968, p. 3). At that time, society's sentiment toward schools was favorable, and the responsibility for augmentation of any deficiencies which may have existed was placed with the family of each individual student (Katz, 1968).

As our country evolved from a primarily agrarian society into an urban industrialized nation, new challenges in educating the masses appeared. Due to an increasing population and a disbanding homogeneous society, our public education system began to lose efficiency and effectiveness. During the first half of the 20<sup>th</sup> century, our nation's Department of Education responded by implementing structural reform that resulted in a "system" which has helped to secure the right for the masses to have a free education (Katz, 1968).

In recent decades the call to reform public education has been the mantra of politicians, private-sector entrepreneurs, parents, and a multitude of others. Understandably, the wide-range of expectations of stakeholders and the fact that schools have become multipurpose institutions expected to address every social, economic and political ill ensure that some aspect of society will always be unhappy with public

education (Reese, 2007). Even so, schools remained critical entities and were indispensable to a republic (Tyack, 2000).

### Sputnik Crisis

The launch of Sputnik by the Soviet Union in the late 1950s, raised concern with the American public that the Soviet school system was superior and would produce superior scientists. Consequently, American public schools received seemingly endless negative press regarding their ability to compete on an international level (Reese, 2007). Press reports indicated public schools were not preparing children in the areas of mathematics and science and insinuated that our nation's education system was in need of a serious overhaul. The United States Office of Education's "life adjustment education" curriculum, which sought to make school relevant to teenagers, was picked apart. Speeches and interviews given by politicians, military leaders and scientists called for a return to a more traditional form of education. Sputnik set a nasty precedent in which public education received all the blame for America's social crisis; yet, when the crisis was resolved, it received no credit (Bracey, 2007).

In 1959, Vice Admiral Rickover, in his book, *Education and Freedom*, declared: "Our schools are the greatest cultural lag we have today" (Rickover, 1959). As a response to the mounting crisis, Congress poured substantial amounts of money into public education. Accordingly, President Eisenhower signed into law the National Defense Education Act, which bolstered science and math offerings in the school setting (Mondale & Patton, 2001).

## Elementary and Secondary Act of 1965

American schools were under fire again in the early 1960s as dissatisfaction with schools was increased. For the first time, a school accountability movement, due to the large amount of public funds being allocated, was propagated (Bowers, 1991). Public perception of schools had never fully recovered from the Sputnik Crisis, and newly designed large standardized tests such as the National Assessment of Educational Progress (NAEP) showed students still lacked basic skills (Vlaanderen, R. as cited in Bowers). Additionally in the 1960s, the Civil Rights Movement had been roiling under the surface due to racial inequalities, including those in the area of education.

In 1964, Congress enacted the Civil Rights Act, which banned racial discrimination in all federally funded institutions and gave the United States Department of Education the authority to collect racial data from schools (Brown, 2004). Pressure mounted to increase educational opportunities for all children, and the Elementary and Secondary Act of 1965 (ESEA) was established. At the time, ESEA was considered the most monumental effort the federal government had attempted in an effort to influence state and local boards to reform their practices (Kantor, 1991). Significant monies were included to entice states and school districts to align to ESEA policies. Over four billion dollars in aid (ie, Title I funds) was offered to assist disadvantaged youth in the educational process (Mondale & Patton, 2001). In addition to assisting disadvantaged youth, funding was also set aside to assist in the areas of education technology, professional development, class-size reduction, safe and drug free schools, bilingual

education, Native American education, charter schools and head start (Electronic Summary of The Elementary and Secondary Education Act of 1965, 2008).

A major expectation of the ESEA was that schools using governmental funds were expected to improve student academic performance of at-risk students in the areas of reading and math, when compared to other students who were not at risk, throughout their state (Wong & Meyer, 1998). In addition, the ESEA was to be evaluated every four years to insure goals were met (Electronic Summary of The Elementary and Secondary Act of 1965, 2008).

#### A Nation at Risk

In 1969, the Gallup organization began to influence national opinion of public education by producing surveys regarding education. These surveys, while largely reflections of public opinion, brought further criticism to school systems. Gallop survey data showed that in 1978, 41 percent of respondents felt public schools were not as good as they used to be, while only 35 percent felt they had improved (Tyack & Cuban, 1995).

With national opinion of public education again suggesting students were not as well educated as their parents; an effort to reform education was instituted by Congress. In 1981, Secretary of Education T. H. Bell was directed by Ronald Reagan to create the National Commission on Excellence in Education, which linked our country's poor economic conditions to a failing public education system (Tyack & Cuban, 1995, p. 14). The commission was directed to pay special attention to teenage youth while critiquing our public schools and universities (Electronic Summary of *A Nation at Risk*, 2001).

In addition to the commission being accountable for teenage youth, it was directed to report on the following: the status of teaching and learning in our nation's public and private schools, colleges, and universities; the performance of American schools and colleges in relationship to other advanced nations; studying student achievement in high school as it relates to college admissions; identifying educational programs which result in notable student success in college; assessing how major social and educational changes in the last quarter century have impacted student achievement; and identifying and defining future considerations that need to be overcome to achieve excellence in education (Electronic Summary of a *Nation at Risk*, 2001).

*A Nation at Risk* brought to light the considerable decline of our nation's school system. In the report, the commission reiterated that the positive gains made from reform achieved after the Sputnik crisis had all but been diminished. Additionally, the commission stated that multiple demands that other social, personal and political institutions cannot tackle or solve continue to be placed on the nation's schools and universities and that these demands detract from the purpose of the educational system (Electronic Summary of *A Nation at Risk*, 2001).

Because of America's declining economic status and the thriving economies of our old World War II foes, angst about the perceived decline in public education resulted in President Ronald Reagan calling for reform. Reagan suggested that not only was the problem economic but also that the enforcement of civil rights legislation contributed to the decline in basic education. Many educators reputed these claims (Mondale & Patton, 2001). Author Nicholas Lemann supported the view of these educators (as cited in



Mondale & Patton), stating that NAEP scores showed the decline was not as dramatic as reported in *A Nation at Risk*. Although public education had many proponents, the quality of America's public schools continued to be questioned by the public.

#### Improving America's Schools Act and Goals 2000: Educate America Act

The year 1994 marked a time in which the federal government embarked on a mission to stimulate a standards-based school reform movement at the state and local levels. Although the federal government set into law three legislated acts, two – The Improving America's Schools Act and Goals 2000: Educate America Act – focused on reauthorizing current education policy and aimed at establishing National Education Goals. The underlying theme of both legislated pieces was the premise that all students should be immersed in curriculum set in high standards instead of educators relying on remedial programs as a means of educating students (United States Department of Education, 1999).

The reauthorization of ESEA in 1994, known as the Improving America's Schools Act (IASA) of 1994, continued the federal government's role as a promoter of equity and excellence by influencing state and local school systems to accept new reform movements through grants and entitlements. By promoting a new policy framework, states and Local Education Agencies (LEAs) were provided funds with which to create high standards and align curriculum. Furthermore, they were given flexibility in how they will create local-based initiatives and were encouraged to create partnerships between schools, parents and communities. The emphasis on teaching and learning as well as the

encouragement given to states and LEAs to be responsible for higher student achievement allowed the federal government to initialize the standards-based reform movement (United States Department of Education, 1999).

IASA introduced the concept of Adequate Yearly Progress (AYP). Under Title I of IASA, the federal government defined AYP as having students, particularly economically disadvantaged and Limited English Proficient (LEP) students, demonstrate “continuous and substantial” yearly progress by meeting the state’s proficient and advanced levels of performance. It provided that each state was to be given the authority to determine the criteria of which it would gauge AYP, and it also provided that any LEA that received funds from the federal government would be required to adhere to the AYP standards established by the state or risk losing the funds (United States Department of Education, 1994a).

In addition to defining AYP, Title I of IASA required states to establish challenging performance standards and to design and implement valid state assessments to measure student performance against the standards. It also held states and LEAs accountable for the success of all students. By incorporating these provisions into IASA the federal government solidified the standards-based reform movement (Goertz, 2001).

By establishing Goals 2000: Educate America Act (Goals 2000), Congress continued to demonstrate the federal government’s desire to reform education at the state level by setting National Education Goals. Title I of Goals 2000 sets National Education Goals by stating that by the year 2000: all children will start school ready to learn; the graduation rate of all schools will increase to 90 percent; all children leaving the 4<sup>th</sup>, 8<sup>th</sup>

and 12<sup>th</sup> grades will demonstrate competency over challenging subject matter and will use their minds well so they can be good citizens and productive in the work force; all teachers will be provided opportunities to improve their skills so they can affect student success in the 21<sup>st</sup> century; the United States will be rated number 1 in the world in terms of student achievement in the areas of science and mathematics; every adult will be literate and will possess the skills necessary to be successful in the global work place; every school will be free of drugs and violence and will provide an environment conducive to learning; parents and schools will conduct partnerships to improve the educational, social and emotional conditions of each school (United States Department of Education 1994b).

Since its inception Goals 2000 was mired in controversy. Liberal members of Congress felt it was proposed as a way to introduce high-stakes testing that would promulgate the lack of student success in problem-plagued schools. Conservative members of Congress felt it was an attempt to supplant local authority of school boards and claimed that it moved away from a back to basics approach by making schools responsible for social issues (Hoff, 1998).

Germane to the legislation and foreshadowing the direction of future National Education Policies was the federal government's desire to advance higher standards in state curriculums. These standards were to be developed on the basis of what children should know and be able to do. Additionally, school improvement strategies were to focus on how to improve student achievement (United States Department of Education, 1998).

Grant funds, distributed on a competitive basis, were dispersed to LEAs through state awards. By attaching large sums of money (1.7 billion over a 4 year period) to the policy, the federal government promoted education reform at the local level. At least 90% of the allocated funds were given to LEAs to improve staff development and encourage local school reform (United States Department of Education, 1998).

Standards-based reform was the new mantra of the federal government. It was believed that if schools incorporated rigorous standards, students would universally improve their performance. Educators on the other hand expressed apprehension that standards-based reform was too closely related to business models (Ohanian, 2000).

Lingering questions still remained once Goals 2000 was implemented. In April 1997, The California Academic Standards Commission of the State Board of Education stated that teachers who understood state standards should be able to prepare and deliver students to the next grade so they can excel in the next set of standards (Ohanian, 2000). Questions arose about how standards-based reform was being actuated in the classroom, and concerns were introduced about how much teachers understood about the new standards being promoted in their states and whether curriculum was actually being aligned to the standards (United States Department of Education, 1994b).

### Standardized Testing

Assessing student performance and the performance of teachers was a driving force behind educational reform movements since the 1970s. Policymakers learned that test results can be used as a vehicle to implement educational reform (Madaus, 1985).

Consequently, standardized tests, due to their relatively inexpensive implementation, their ability to be easily mandated, their flexibility and ease of implementation, and their transparency, have provided a mechanism through which reformers rallied support in an effort to overhaul our nation's public education system (Linn, 2000).

Standardized tests came into existence just before the turn of the twentieth century. Initially, standardized tests were not used at the national level to compare student performance. Instead, they were used as means to compare teacher performance within schools, between schools, and between districts and to see how students performed on outside exams (Resnick, 1980a).

By World War I, test designers began to provide data that enabled educators to make comparisons on a more global scale. They developed exams that presented data about representative populations and allowed students to be placed into homogeneous groups. Although data was not always used effectively, the exams did assure they were implemented with some form of uniformity by providing directions for conformity of implementation and interpretation and by norming grading criteria (Resnick, 1980).

The role of standardized testing was diminished from 1930 through the 1950s. During this era, Americans had confidence in the public education system (Resnick, 1980). In the 1960s, politicians and reformists called for all children to have equal access and equal educational opportunities in our nation's public education system. A joint committee of legislators was formed to promote these ideals, and the Elementary and Secondary Education Act of 1965 (ESEA) was established to provide financial assistance to schools and districts with substantial numbers of low-income students. In addition, the

federal government contracted with the Education Commission on States and created the National Assessment of Educational Progress (NAEP) in 1969. The goal of the establishment of NAEP was to monitor achievement in 10 learning areas and to assess change in achievement over the years. Information gathered from NAEP was to then be reviewed and used by policymakers to institute change in public education (Resnick, 1980).

ESEA funded a variety of programs to facilitate widespread reform in the public education system. ESEA consisted of six titles. The most significantly funded of these was Title I, which held a provision for required testing of student performance as a means of evaluating funded programs (Resnick, 1980). The significant funding led to a boon in the testing industry, and norm-referenced testing with Normal Curve Equivalents (NCE) was introduced. Essentially, NCEs were standardized with a mean score of 50 and a standard deviation of 21.06 (Linn, 2000).

To accomplish the implementation of NCEs, the Title I Evaluation and Reporting System (TIERS) was established, and testing companies experienced large-scale financial success (Tallmadge & Wood, 1981). TIERS data were collected and information presented to the general public painted a relatively positive picture. Elizabeth Reisner and her colleagues, who were commissioned to study TIERS in 1982, reported: 1) the program provided the achievement data requested by Congress as specified, 2) the technical assistance and evaluation model had a positive effect on instructional practices at the local level, and 3) new approaches to the evaluation process were introduced to educators (Reisner, et. al., 1982).

While TIERS appeared to be successful in evaluating and reporting NCEs, other studies produced different opinions about the success of TIERS. Linn and his colleagues found a number of variables such as student selection, scale conversion errors, administration conditions and teaching to the test to be among factors which contributed to potential corruption of the data and inflated results (Linn, et al. 2002)

Minimum Competency Exams (MCTs), as a means for large scale standardized testing, became the norm in the 1970s and 1980s. States, such as Florida, were closely studied due to their requiring students to pass their version of an MCT, the High School Competency Test (HSCT), in order to graduate (Linn, 2000). According to Linn (2000), at the inception of the HSCT in 1977, white and Hispanic students performed well on their first attempt to pass. African-American students performed considerably lower, but their scores did increase as the test was implemented over time (Linn, 2000). The underlying lesson learned from the long-term implementation of the HSCT and other MCTs was that student performance becomes flat over time. Straight-lined standardized test data over long periods of time should be reviewed with caution when using it to drive policy (Linn, 2000).

In the 1990s, the promotion of higher standards by federal and state legislatures resulted in a standards-based reform movement and the development of standards-based assessments. Swanson and Stevenson (2002), in their article, “Standards-Based Reform in Practice: Evidence on State Policy and Classroom Instruction from the NAEP State Assessment” describe school reform efforts in the 1970s and 1980s as largely unsuccessful due to their lack of connection between national policy and the classroom.

On the other hand, they state that standards-based reform, when effective standards-based assessments were included, can lead to effect improvement of instructional practices (Swanson & Stevenson).

According to Blum (2000), instituting higher content standards and establishing a standards-based assessment system allowed states to set benchmarks and performance standards for student achievement. By the mid-to-late 1990s, almost every state in the nation had adopted some form of standards-based assessment (Blum, 2000).

#### The National Assessment of Educational Progress

The National Assessment of Education Progress (NAEP) was in existence since 1969. It was established by the United States Department of Education as a means to compare student performance data in the areas of mathematics, reading, writing and science. The assessment was conducted in states by randomly selecting fourth, eighth, and twelfth graders to sit for subject area exams (United States Department of Education, 2008b).

The National Center for Educational Statistics (NCES) was responsible for developing, administering and scoring the test. NCES fell under the jurisdiction of the United States Department of Education's Institute of Education Services (IES) and was given the responsibility of analyzing NAEP data and assuring it was accessible (United States Department of Education, 2008b).

A governing board of 26 people, known as the National Assessment Governing Board (NAGB), was responsible for determining the content of the test. The board



included educators, governors, legislators, business owners, and lay people, all of which were appointed by the United States Secretary of Education. The NAGB played a key role in establishing NAEP Frameworks which guided test developers in constructing the exam. NAEP Frameworks were content standards and determined what students should know and be able to do at the grade level being assessed. Once NAEP Frameworks were established, test questions were developed, scrutinized and field tested before becoming a part of the published exam (United States Department of Education, 2008b).

Students were selected to participate in NAEP using a sampling procedure designed to collect data from representative populations of students in the United States. NAEP randomly selected students so data based upon racial, ethnic, geographical and socioeconomic diversity was collected. Approximately 3000 students in each state were selected for each grade and subject that was being tested (United States Department of Education, 2008b).

Questions on NAEP were presented in both multiple choice and open-ended formats. The integrity of the data collected by NAEP was critical in its ability to be accurately analyzed and reported. Computers were used to optically scan and grade multiple choice questions. On the open-ended sections of NAEP, scorers were hired based on qualifications and experience. Manuals, trainings and reviews of each scorer's performance were all incorporated to assure quality and consistency of scoring (United States Department of Education, 2008b).

NAEP results were reported in scales and by achievement levels. Scales ranged from 0-500 and were reported as average scores of a group of students. Achievement

levels were set by the NAGB and reflected what students should know and be able to do in the subject and at the grade level that was being tested. Achievement levels were reported by the percentages of students performing at the Basic, Proficient or Advanced levels (National Assessment Governing Board, 2008). A student who scored at the Basic Level demonstrated minimal knowledge of the subject matter being tested. If a student scored at the Proficient Level he or she has demonstrated mastery over the challenging content of the subject matter. A student who scored at the Advanced level demonstrates superior performance on the specific exam (United States Department of Education, 2008b).

In 2002, Congress reauthorized NAEP and mandated it provide a fair and accurate measurement in reading and other subjects. Congress expected NAEP to report trends in student reading proficiency at both state and national levels. In addition, NAEP served as a means of informing policymakers about relationships between student performance and significant background variables as well as a method of providing comparative data on the performance of groups, states, and regions (National Assessment Governing Board, 2008).

A fundamental goal of NAEP was to provide accurate, consistent student achievement data so educators, researchers, policy-makers, and the general public can make informed decisions about the progress of education in America. By becoming the “Nation’s Report Card,” NAEP was considered the “Gold Standard” in measuring student achievement (Jones & Olkin, 2004).

According to Beaton and Zwick, NAEP was not an ordinary testing program designed to report individual performance of students, schools or school districts. As a result, educational decisions about individual student performance cannot be made (Beaton & Zwick, 1992). NAGB concurred by stating that NAEP's current design did not provide individual student diagnostic information; however, it did provide group performance data without dictating prescriptive actions for Local Education Authorities (LEAs) to incorporate into their plans (National Assessment Governing Board, 2008).

NAEP can play an important role in confirming statistics and in allowing comparisons to be made state by state. In 2002, an Ad Hoc Committee on Confirming Test Results determined that NAEP should be used to confirm state test results by reporting progress of student achievement, evaluating the rigor of state standards and disclosing achievement gaps among significant subgroups of students (Ad Hoc Committee on Confirming Test Results, 2002).

Bert Stoneberg, an NAEP Program Manager from the Office of the Idaho State Board of Education added that the United States Department of Education (USDOE) had not yet provided technical assistance on how to best use NAEP data to confirm state scores. He did, however, state that the USDOE did provide enough information to provide direction when making comparisons. Stoneberg's research suggested comparisons can be made between NAEP's "At or Above Basic" data and a state's Adequate Yearly Progress (AYP) results. He stated that point by point comparisons should not be made. He proposed the use of trend lines, to compare NAEP data and state-created exams, as a reliable method in confirming state AYP results (Stoneberg, 2007).

NAEP did have a practical use in terms of comparing performance of large student populations on a state by state basis. In 1992, the NAGB established cut scores in its scale scores for reading. By establishing cut scores, student performance levels were compared on a national basis (National Assessment Governing Board, 2008). According to Waltman (1997), it would be useful if the NAGP translated its performance standards more broadly so they could have been translated to state assessment scale scores. By aligning with state assessments, the NAGP could eliminate the need for students to take multiple assessments, and monitoring of state-level progress could have been made using the NAEP (Waltman).

Lee pointed out that even though NAEP compared state by state student performance, problems did exist. The tracking of student achievement and the alignment of assessment systems to meet the reporting criteria of NCLB presented challenges for states. Additionally, NAEP was not aligned to specific standards of any given state (Lee, 2007).

#### The No Child Left Behind Act of 2001(NCLB)

On January 8, 2002, President George W. Bush signed Public Law 107-110, The No Child Left Behind Act of 2001 (NCLB), into law. This largely bipartisan effort of Congress was essentially the reauthorization and expansion of the Elementary and Secondary Act of 1965. NCLB represented Congress' effort to bring greater accountability and improvements in student achievement to the forefront in the educational reform movement (Shaul & Ganson, 2005).

The premise of NCLB was set on four pillars: accountability for results, emphasis on doing what works based on scientific research, expanded parental options, and expanded local control and flexibility (United States Department of Education, 2004a). Under NCLB, states were given more local control and were expected to close the achievement gap for all children, including those from economically disadvantaged families. Linn notes, NCLB's emphasis on promoting learning in groups of students who have lagged behind was positive. By focusing on standards based reform and promoting higher state standards, NCLB increased opportunities for students. Conversely, NCLB was driven by accountability and relied too heavily on high-stakes tests that focused primarily on reading and mathematics. In fact, NCLB did not allow states to use other indicators that were not outlined in Sec. 1111 (b)(2)(A)(i) because they could have outweighed how students performed in these areas (Linn, 2003).

Another major provision, Sec. 1111 (b)(1)(A) of NCLB, was the expectation that states set demanding content standards as well as challenging student achievement standards (Linn, 2003). States were required to establish measurement tools to evaluate whether Local Educational Agencies (LEAs) were making AYP within parameters set in Title 1 (United States Department of Education, 2004d). Determining AYP was the means by which a state measured its demanding content and challenging student achievement standards. Each state established its own criteria when assessing a student's yearly progress and reported these scores to parents. The federal government required states to have 100% of their students demonstrate success on state developed standardized tests by the year 2014 (United States Department of Education, 2004d).

States that received federal funds under NCLB were required to have fourth and eighth grade students participate in NAEP reading and mathematics assessments every two years. According to the U.S. Department of Education, “the NAEP provided parents with a wealth of data about the condition of education in the United States” (United States Department of Education, 2004d).

The policies of NCLB were stated in the following ten titles of NCLB: Title I – Improving the Academic Achievement of the Disadvantaged; Title II – Preparing, Training, and Recruiting High Quality Teachers and Principals; Title III – Language Instruction for LEP and Immigrant Students; Title IV – 21<sup>st</sup> Century Schools; Title V – Promoting Informed Parental Choice; Title VI - Flexibility and Accountability; Title VII - Indian, Native Hawaiian, and Alaska Native Education; Title VIII - Impact Aid Program; Title IX - General Provisions; Title X -Repeals, Re-designations, and Amendments to Other Statutes (107<sup>th</sup> Congress, 2002).

In creating NCLB, the federal government’s intent was to influence a change in the way our state public school systems conducted business. Although all of the ten titles were important, for the purpose of the study, the researcher has chose two titles, Title 1 and Title VI, due to their pertinence.

#### Title 1: Improving the Academic Achievement of the Disadvantaged

Title I targets disadvantaged youth by providing federal funds to Local Educational Agencies (LEAs) to ensure equal access to high quality educational opportunities. Its purpose was to “ensure that all children have a fair, equal, and

significant opportunity to obtain a high quality education and reach, at minimum, proficiency on challenging state academic standards and state academic assessments” (United States Department of Education, 2005).

Under NCLB, the accountability provisions of Title I were strengthened in 2001 to provide a mechanism for holding states, school districts, and schools accountable for improving the academic achievement of all students and turning around low-performing schools. Additionally, it required that students in low-performing schools were to be afforded alternatives to enable those students to receive a high-quality education. Schools and districts were to demonstrate all students were meeting annual targets established by states through their assessment systems. Annual assessments were required at grades 3-8 and once in grades 10-12. In addition, states were required to report student performance data by poverty, race/ethnicity, and limited English proficiency. All students were expected to be proficient in reading and mathematics by the 2013- 2014 school year. Schools and districts that do not make Adequate Yearly Progress (AYP) for two consecutive years were subject to increased interventions and must provide additional educational options for students (Institute of Educational Sciences, 2007).

In NCLB, the Adequate Yearly Progress (AYP) provision was strengthened in an effort to focus LEAs on eliminating the achievement gap. By forcing states and LEAs to report AYP by subgroups, Congress could identify where the achievement gaps existed. Autonomy was given to states to determine AYP each year to assess performance of school districts and schools. AYP was intended to demonstrate, through data, where schools needed to improve and where resources should have been focused. Holding

schools accountable for the performance of all students in relationship to state standards was central to the theme of NCLB (Shaul & Ganson, 2005).

Minimal criteria for AYP were clarified as follows:

1. A single statewide accountability system applied to all public schools and LEAs: This criterion assured that all schools were provided the same assessments and were measured by the same accountability system so performance could be reported, transparent and comparable within the state (United States Department of Education, 2005).
2. All public school students were included in state accountability systems: Under this criterion, no less than 95% of students should have been tested. Additionally, student test scores were only considered for AYP if the students were in attendance for a full year at the academic institution. This component minimized the impact of mobile student populations (United States Department of Education, 2005).
3. A state's definition of AYP was based on expectations for growth in student achievement that was continuous and substantial, such that all students were proficient in reading and math no later than 2013-2014: The Federal Government required states to progressively increase statewide proficiency goals that reflected 100 % of students were proficient in reading and mathematics by 2013-2014. Goals were established from 2001-2002 and



should have steadily increased over the 12-year period (United States Department of Education, 2005).

4. A state makes annual decisions about the achievement of all public schools and LEAs: This criterion required states to determine whether a school or district met AYP on an annual basis even if it chose to average AYP scores over a three year period (United States Department of Education, 2005).
5. All public schools and LEAs were held accountable for the achievement of individual subgroups: The federal government defined subgroups as: economically disadvantaged students, major ethnic/racial groups, LEP students, and students with disabilities. In addition, states were required to establish separate annual achievement goals for individual subgroups, but LEAs were given the autonomy to determine how to reach the goals established by the state (United States Department of Education, 2005).
6. A state's definition of AYP was based primarily on the state's academic assessments: By adding this criterion, the federal government required states to use only academic achievement data to determine if a school or LEA was making AYP, thus centralizing the focus on academic improvement (United States Department of Education, 2005).
7. A state's definition of AYP included graduation rates for high schools and an additional indicator selected by the state for middle and elementary schools (such as attendance rates). In this criterion, the federal government defined the graduation rate as "the number of students who graduate from high school in

the standard number of years.” Since middle and elementary schools did not figure graduation rates; they were required to report one additional indicator. When a state selected the other indicator it will report, it was incumbent on the state to select indicators that would not reduce the number of schools that would be subject to school improvement, corrective action or restructuring (United States Department of Education, 2005).

8. AYP was based on separate reading/language arts and mathematics achievement objectives: This criterion mandated all students in identified subgroups should have met 100% proficiency in reading and mathematics by 2013 -2014 (United States Department of Education, 2005).
9. A state’s accountability system was statistically valid and reliable: States were required to maintain reliable and valid statistical systems that took into account the size of each subgroup at each school by determining what the minimal size of each was and what the technical specifications of the tests allowed (United States Department of Education, 2005).
10. In order for a school to make AYP, a state ensured that it assessed at least 95% of students in each subgroup enrolled: States had to assure that all student data was reported in subgroups. In addition, 95% of all students identified in subgroups had to be tested (United States Department of Education, 2005).

Fundamental to Title I was closing the achievement gap between low and high performing students who were minority and nonminority students as well as those who

were disadvantaged and advantaged students. LEAs were expected to address achievement gap disparities by using data as a means of improving the instruction and learning of the disadvantaged.

The appropriation of funds to accomplish the goals of Title I was paramount to its success. In 2007, Congress authorized the allocation of 25 billion dollars to meet the needs of our nation's highest poverty schools in an effort to target financial resources to the districts and schools where needs were greatest (United States Department of Education, 2004b). Funding was used to provide additional instructional staff, professional development, extended-time programs, and other strategies for raising student achievement in high-poverty schools. While increased discretion and flexibility was given to LEAs in using provided financial resources, the expectation was that they were held more accountable in increasing student performance (The United States Department of Education, 2004d).

States accepted the terms of NCLB because it was financially beneficial. Most important was the setting of appropriate performance standards and cut scores. Because guidance from the federal government on how to set performance standards and cut scores was minimal, it was logical that student achievement comparisons from state to state was not sensible (Schafer, et al., 2007). The lack of being able to make sensible state by state comparisons was a reason that a national assessment system was considered.

## Title VI: Flexibility and Accountability

The principal intent of creating Title VI, by the federal government, was to create collaborative environments in which states, LEAs and higher institutions developed assessments in order to elevate student performance and close the achievement gap. Money provided in this area could be combined or transferred to allow more flexibility. In return for providing flexible opportunities with the funds, the federal government held states and local LEAs highly accountable for improving student achievement (United States Department of Education, 2007b).

Section 6111 of Sub Part 1 of the Accountability section of Title VI housed the majority of the provisions regarding the federal government's desire to influence states to align their accountability systems to academic content and standards. States were provided funding in the form of grants, to aid them in establishing accountability systems or to enhance their existing assessment systems (United States Department of Education, 2004c). Formula grants were provided and termed as "Grants for State Assessment and Related Activities." States could use formula grants to develop additional academic standards and administer assessments as required under NCLB. Each state received a minimum of 3 million dollars plus an additional amount of money which was determined by the number of children who resided in the state between the ages of 5-17. Flexibility was given to states to use the grants, if high academic standards were in place and quality assessments had been developed, to ensure LEAs were held accountable for results (United States Department of Education, 2007b).

A vehicle for states and LEAs to compete for grants was also established under Title VI in the form of competitive grants known as “Grants for Enhanced Assessment Instruments.” Competitive grants were available after the distribution of the formula grants, if funds were available, and supported collaborative efforts of states and institutions to enhance the reliability and validity of their assessment instruments. These funds were allocated based on a state’s need and the quality and scope of their application (United States Department of Education, 2007b). By influencing states to assess academic content and standards, LEAs were expected to align their curriculum with state standards. Additionally, the federal government provided additional grant funds for districts to improve curriculum alignment and to ensure instructional materials were available (United States Department of Education, 2004c).

Further, Section 6111 of Sub Part 1 of the Accountability section of Title VI addressed the need for states to become more transparent with collected data by forcing states to develop information reporting systems for parents and communities. Information presented was related to student achievement. Strategies developed from the information gleaned by educators was scientifically-based and reflect best practices. By creating transparency of the data and putting it in the hands of LEAs, appropriate staff-development was established to positively affect student achievement (United States Department of Education, 2004c).

The validity and reliability of accountability measures addressed in Title VI were central to federal government’s desired outcomes. In an effort to bolster state assessment systems grants and assistance opportunities were made available to states (United States

Department of Education, 2004c). It was important to note that states still have the flexibility and authority to determine what standards were considered important for their citizens and how those standards were assessed and measured.

Linn, Baker and Bettebenner (2002) assert that the challenge for states will be to honor the intent of Title VI by providing information that was scientifically-based and that can improve instruction. It was important that information provided by state assessments was relevant and utilized by all stakeholders. Under Title VI, states and LEAs were challenged with producing critical analysis of data so they made the most of it in their systems. States that invested in long-term studies on how their assessments were implemented and on the impact of the results of their implementation had the greatest opportunity to impact student achievement (Linn, et al.).

Subpart 2 of Title VI, known as the State and Local Transferability Act, enhanced States and LEAs flexibility and allowed states to combine resources to meet the needs of the citizens. This critical piece of legislature freed up Title I funds and allowed those funds to be put to use where they were most needed. Allowing states and LEAs to transfer the large funds attached to Title I created additional opportunities for them (United States Department of Education, 2004

In Title VI, Subpart 3 - State and Local Flexibility, - additional provisions were stated to address the intent of the federal government. Under Section 6132, states and LEAs were given greater authority and freedom in determining how to improve student achievement and to employ school reform measures. It was believed that by empowering administrators, teachers, parents and students in the educational process, overall student

achievement would improve. Removing barriers to local reform efforts, along with providing resources and the flexibility to use the reform efforts, demonstrated the federal government's intent to provide high quality educational opportunities for all children and its desire to close the achievement gap (United States Department of Education, 2004c).

### Summary

In the 1960s a school accountability movement, due to the large amount of public funds being allocated, was propagated (Bowers, 1991). A joint committee of legislators was formed, and the Elementary and Secondary Education Act of 1965 (ESEA) was established to provide financial assistance to schools and districts with substantial numbers of low-income students. As a result, ESEA became the largest effort of the federal government to influence state and local boards to reform their practices (Kantor, 1991).

In addition, the federal government contracted with the Education Commission on States and created the National Assessment of Educational Progress (NAEP) in 1969. The goal of the establishment of NAEP was to monitor achievement in 10 learning areas and to assess change in achievement over the years. Information gathered from NAEP was to then be reviewed and used by policymakers to institute change in public education (Resnick, 1980). By becoming the "Nation's Report Card," NAEP was considered the "Gold Standard" in measuring student achievement (Jones & Olkin, 2004).

In the 1970s assessing student performance and the performance of teachers became a driving force behind educational reform movements. Policymakers, by using

standardized tests as a means to implement policy, learned that test results could be used as a vehicle to implement policy (Madaus, 1985). Consequently, standardized tests, due to their being relatively inexpensive to implement, their ease to mandate, their flexibility and ease of implementation, and their transparency, have provided a mechanism for reformists to rally support as a means of overhauling our nation's public education system (Linn, 2000).

In 1981, the National Commission on Excellence in Education, which linked our country's poor economic conditions to a failing public education system, was commissioned (Tyack & Cuban, 1995). The commission was directed to report the status of teaching and learning in our nation's public and private schools as well as in other areas considered important by Congress (Electronic Summary of *A Nation at Risk*, 2001). The commission's report, *A Nation at Risk*, brought to light the considerable decline of our nation's school system (Electronic Summary of *A Nation at Risk*, 2001). As a result of this report, President Ronald Reagan called for reform (Mondale & Patton, 2001).

In the 1990s, the federal government promoted education reform by embarking on a mission to stimulate a standards-based school reform movement at the state and local levels. Subsequently, the federal government set into law two legislated acts – The Improving Americas Schools Act and Goals 2000: Educate America Act – that focused on reauthorizing current education policy and aimed at establishing National Education Goals. The underlying theme of both legislated pieces was the premise that all children should be immersed in curriculum containing high standards instead of relying on



remedial programs as a means of education (United States Department of Education, 1999).

Additionally, in 1992 the National Assessment Governing Board (NAGB) established cut scores in its scale scores for reading on NAEP. By using established cut scores, student performance levels were compared on a national basis (National Assessment Governing Board, 2008). NAEP was conducted in states by randomly selecting fourth, eighth, and twelfth grade students to sit for subject area exams (United States Department of Education, 2008b).

On January 8, 2002, President George W. Bush signed Public Law 107-110, The No Child Left Behind Act of 2001 (NCLB), into law. Under NCLB, the accountability provisions of Title I were strengthened to provide a mechanism for holding states, school districts, and individual schools accountable for improving the academic achievement of all students (Institute of Educational Sciences, 2007). A major provision of NCLB was the expectation that states set demanding content standards as well as challenging student achievement standards (Sec. 1111(b)(1)(A), as cited in (Linn, 2003). Determining AYP was the means by which a state measured its demanding content and challenging student achievement standards. Each state established its own criteria for assessing a student's yearly progress and was required to report scores to parents. The federal government required states to have 100% of their students demonstrate success on state developed standardized tests by the year 2014 (United States Department of Education, 2004d).

States accepted the terms of NCLB because it was financially beneficial. States that received federal funds under NCLB were required to have fourth, eighth and twelfth

grade students participate in NAEP reading and mathematics assessments every two years (United States Department of Education, 2004d). Most important was the setting of appropriate performance standards and cut scores by states. Because guidance from the federal government on how to set performance standards and cut scores was minimal, it was logical that student achievement comparisons from state to state was not sensible (Schafer et al., 2007). The lack being able to make sensible state by state comparisons was a reason that a national assessment system needed to be considered.

In 2002, Congress reauthorized NAEP and mandated that it was fair to provide a fair and accurate measurement in reading and other subjects. Congress expected NAEP to report trends in student reading proficiency at both state and national levels. In addition, NAEP served as a means of informing policymakers about the relationship between student performance and any significant background variables as well as a method of providing comparative data on performance of groups, states, and regions (National Assessment Governing Board, 2008). Additionally in 2002, an Ad Hoc Committee on Confirming Test Results reported that “differences between the National Assessment of Educational Progress and state testing programs must be explored and reported” (Stoneberg, 2007).

## CHAPTER THREE: METHODOLOGY

### Introduction

The intent of this study was to provide information for policymakers and practitioners by comparing performance of eighth grade students in 2007 on state standardized reading assessments and by the four census-defined regions identified by NAEP. As highlighted in the review of literature, the Ad Hoc Committee on Confirming Test Results (2002) reported that “differences between the National Assessment of Educational Progress and state testing programs must be explored and reported” (Stoneberg, 2007).

In 2007, the National Center for Educational Statistics, a division of the United States Department of Education, administered and scored the NAEP exam in reading by randomly selecting eighth grade students in each of the 50 states to achieve a representative population of the country (United States Department of Education, 2008b). The data were reported as part of the Nation’s Report Card.

Sec. 1111(b) (1) (A) of NCLB required states to set demanding content standards as well as challenging student achievement standards (Linn, 2003). States were required to establish measurement tools to evaluate whether Local Educational Agencies (LEAs) were making progress with all students (United States Department of Education, 2004d). In 2007, all 50 of the United States conducted eighth grade reading assessments and reported data related to the tests.

## Statement of the Problem

The purpose of this study was to compare performance of eighth grade students in 2007 on state standardized reading tests and by the four census-defined regions identified by NAEP. The researcher chose to compare all states due to a NCLB requirement that all states report eighth grade student reading proficiency data and the availability of NAEP scores of eighth grade students by state. Furthermore, this study attempted to show that a national assessment system provided more accurate comparisons between states to validate the provisions of NCLB.

## Research Questions

The study was guided by the following research questions:

1. What relationship exists, if any, between the percent of eighth grade students identified as proficient in reading on state assessments and on the NAEP in 2007?
2. What comparison can be made, if any, between the percent of eighth grade students identified as proficient and above in reading on state assessments and on the NAEP, in the four census-defined regions identified by the NAEP in 2007?
3. What relationship exists, if any, between the percent of low socioeconomic eighth grade students identified as proficient and above in reading on state assessments and on the NAEP, in the four census-defined regions identified by the NAEP in 2007?
4. What relationship exists, if any, between the percent of nonwhite eighth grade students identified as proficient and above in reading on state assessments and on the NAEP, in the four census-defined regions identified by NAEP in 2007?

## Population and Sample

Eighth grade student reading performance data on state assessments and the NAEP in 2007 were targeted for this study. As a part of the requirements of NCLB, states were required to report eighth grade reading performance data and were required to have students participate in the NAEP. NAEP scores were reported by the United States Department of Education. State scores were available for constituents of the state and were reported to the United States Department of Education.

## Instrumentation

The researcher collected NAEP data from the National Center for Education Statistics, a well respected division of the United States Department of Education's Institute of Education Services. State performance data were collected from the United States Department of Education SY 2006-2007 Consolidated State Performance Report. Data collected were analyzed through a series of statistical procedures by the researcher for the purpose of the study.

## Instrument Reliability and Validity

### Background

The federal government contracted with the Education Commission on States and created the National Assessment of Educational Progress (NAEP) in 1969. Their goal in establishing the NAEP was to monitor achievement in 10 learning areas and to assess change in achievement over the years. Information gathered from NAEP was to then be

reviewed and used by policymakers to institute change in public education (Resnick, 1980).

The accountability provisions of Title I were strengthened under NCLB to provide a mechanism for holding states, school districts, and schools accountable for improving the academic achievement of all students (Institute of Educational Sciences, 2007). A major provision of NCLB was the expectation that states set demanding content standards as well as challenging student achievement standards (Sec. 1111(b)(1)(A), as cited in (Linn, 2003). Determining Adequate Yearly Progress (AYP) was the means by which a state measured its demanding content and challenging student achievement standards. Each state established its own criteria for assessing a student's yearly progress and was required to report these scores to parents. The federal government required states to have 100% of their students demonstrate success on state-developed standardized tests by the year 2014 (United States Department of Education, 2004d).

### Reliability

According to Jones and Olkin (2004), NAEP was considered the "Gold Standard" in measuring student achievement. The Education Testing Service (ETS) was responsible for the design, analysis and reporting of the 2007 NAEP reading data. NAEP assessments contain both multiple-choice and constructed response questions. Constructed response questions required students to provide their own answers. In order to score large numbers of constructed response questions, over 3 million annually, with a high level of accuracy, NAEP incorporated extensive quality control measures to ensure reliability (United States

Department of Education, 2008a). It was incumbent on each state to develop its own procedures to assure accurate and reliable design, analysis and reporting of state assessments under the terms of NCLB.

### Data Collection

In November 2008, eighth grade reading performance data were collected from the NCES 2007 State Snapshot Report (Appendix A) and from SY 2006-2007 Consolidated State Performance Reports (Appendix B). The data were organized in an SPSS worksheet and disaggregated by the four census-defined regions identified by NAEP (Figure 1). Data were further disaggregated by low socioeconomic students and nonwhite students. This information was analyzed using several statistical procedures to answer the four research questions.

### Data Analysis

The following 2007 demographic data were collected within an SPSS worksheet: (a) name of state, (b) the four census-defined regions of NAEP, (c) the percent of eighth grade students identified as proficient and above in reading on NAEP, (d) the percent of low socioeconomic eighth grade students identified as proficient and above in reading on state assessments, (e) the percent of low socioeconomic eighth grade students identified as proficient and above on NAEP, (f) the percent of nonwhite eighth grade students identified as proficient and above in reading on state assessments, and (g) the percent of nonwhite eighth grade students identified as proficient and above in reading on NAEP.

## Summary

Through the analysis of data collected, the researcher was able to compare eighth grade students in reading in 2007 on the NAEP and state standardized tests. Additionally, data was disaggregated by census-defined regions of the NAEP and compared. Once defined into regions, data was further evaluated between low socioeconomic students on the NAEP and state standardized tests. Finally, data was further compared between nonwhite students on the NAEP and state standardized tests. Through the answering of the four research questions, policymakers and stakeholders will be able to make more accurate comparisons of students' performance to validate the provisions of NCLB.



## CHAPTER FOUR: ANALYSIS OF DATA

### Introduction

This investigation offered insight into the testing provisions of No Child Left Behind (NCLB) and provided information for policymakers and practitioners by comparing performance of eighth grade students in 2007 on state standardized reading assessments and by the four census-defined regions identified by NAEP. According to Shaul and Ganson (2005), the formation of the NCLB was an effort on the part of Congress to bring greater accountability and improvements in student achievement to the forefront in the educational reform movement (Shaul & Ganson). Due to an increased transparency of student performance data and the lack of uniform accountability systems, the inability of educators, policy-makers and the general public to make state-by-state comparisons in the area of reading was apparent after NCLB (National Center for Educational Statistics, 2007).

To keep the enormity of the NCLB's policies manageable, two titles were chosen to guide the analysis of data for the research: Title I - Improving the Academic Achievement of the Disadvantaged and Title VI - Flexibility and Accountability. The main purpose of Title I was to "ensure that all children have a fair, equal, and significant opportunity to obtain a high quality education and reach, at minimum, proficiency on challenging state academic standards and state assessments" (United States Department of Education, 2005). The principal goal of Title VI was to create flexible environments in which states, Local Education Agencies (LEAs) and higher institutions developed

assessments in order to elevate student performance and close the achievement gap (United States Department of Education, 2007b).

The intent of this study was to provide information for policymakers and practitioners by comparing performance of eighth grader students in 2007 on state standardized reading assessments and by the four census-defined regions identified by National Assessment of Educational Progress (NAEP). As highlighted in the review of literature, the Ad Hoc Committee on Confirming Test Results (2002) reported that “differences between the National Assessment of Educational Progress and state testing programs must be explored and reported” (Stoneberg, 2007).

In 2007, the National Center for Educational Statistics (NCES), a division of The United States Department of Education (USDOE), administered and scored the NAEP exam in reading by randomly selecting eighth grade students in each of the 50 states to achieve a representative population of the country (United States Department of Education, 2008b). The data were reported as part of the Nation’s Report Card.

Sec. 1111(b) (1) (A) of NCLB required states to set demanding content standards as well as challenging student achievement standards (Linn, 2003). States were required to establish measurement tools to evaluate whether LEAs were making progress with all students (United States Department of Education, 2004d). In 2007, each of the 50 United States conducted eighth grade reading assessments and reported data related to the tests.

## Purpose of the Study

Specifically, the researcher designed the study to provide comparisons of the proficiency of eighth grade students in reading on the NAEP and state standardized tests in 2007. Additionally, data were disaggregated into the census-defined regions of the NAEP and then compared. Once divided into census-defined regions, data between low socioeconomic students on the NAEP and low-socioeconomic students on state standardized tests were further evaluated. Additionally, the scores of nonwhite students on the NAEP and scores of nonwhite students on state standardized tests were compared. Due to the NCLB requirement that all states report eighth grade reading proficiency scores and the availability of NAEP scores of eighth grade students by state, the researcher was able to compare scores from all states.

## Research Questions

The study was guided by the following research questions:

1. What relationship exists, if any, between the percent of eighth grade students identified as proficient and above in reading on state assessments and the NAEP in 2007?
2. What comparison can be made, if any, between the percent of eighth grade students identified as proficient and above in reading on state assessments and on the NAEP, in the four census-defined regions identified by the NAEP in 2007?
3. What relationship exists, if any, between the percent of low socioeconomic eighth grade students identified as proficient and above in reading on state assessments and on the NAEP, in the four census-defined regions identified by the NAEP in 2007?
4. What relationship exists, if any, between the percent of nonwhite eighth grade students identified as proficient and above in reading on state assessments and on the NAEP, in the four census-defined regions identified by NAEP in 2007?

## Demographics

Data were collected from two sources. State data were gathered from the United States Department of Education's SY 2006-2007 Consolidated State Performance Report (2006-2007 CSPR). NAEP data were collected from the National Center for Education Statistics, a division of the United States Department of Education's Institute of Education Services. From the research, eight variables were deemed important to answer the four research questions. The variables were: a) States, b) Census-Defined Regions of the NAEP, c) State Proficient and Above, d) NAEP Proficient and Above, e) Low socioeconomic State, f) Low socioeconomic NAEP, g) Nonwhite state and h) Nonwhite NAEP.

An example of the data sources were: a) The National Center for Educational Statistics State Snapshot Report, found in Appendix A, b), The Consolidated State Performance Report for the State of Alabama's 2007 eighth grade students' reading performance found in Appendix B, and c) The National Center for Education Statistic's Four Census-Defined Regions of NAEP, found in Appendix C.

The researcher used the collected information to establish an SPSS worksheet for statistical calculations. All categorical information was derived directly from collected data with the exception of nonwhite state and nonwhite NAEP. In order to make accurate comparisons, the researcher used mathematical procedures to establish accurate data for two categories. To establish a percentage of nonwhite student data from the 2006-2007 CSPR, the number of nonwhite students scoring at or above the proficient level was divided into the number of all students who completed the assessment and for whom a

proficiency level was assigned. To establish a percentage of nonwhite data from the NAEP, the percentage of students from each nonwhite category who sat for the exam was multiplied by the percentage of nonwhite students scoring at or above proficiency. The raw data that were established in an SPSS worksheet for the purposes of the research are presented in Appendix D.

### Analysis of Data

This section was arranged according to the four research questions that guided this study. The research questions were stated, followed by a discussion of the data.

#### Research Question 1

What relationship exists, if any, between the percent of eighth grade students identified as proficient and above in reading on state assessments and the NAEP in 2007?

The desired outcome was to gain insight as to whether there was a relationship between eighth grade student performance on state reading assessments in 2007 and eighth grade student reading performance on the NAEP in 2007. Data were collected from the 2006-2007 CSPR and from the NAEP State Snapshot Report on each individual state and compiled into an SPSS spreadsheet, as shown in Appendix E.

The researcher initially established a scatter plot, which was useful in determining whether a linear relationship was present. Statistical assumptions were visually examined using the scatter plot. Based on the scatter plot, the state proficient and above and NAEP proficient and above data had a significant relationship, as displayed in Figure 2.

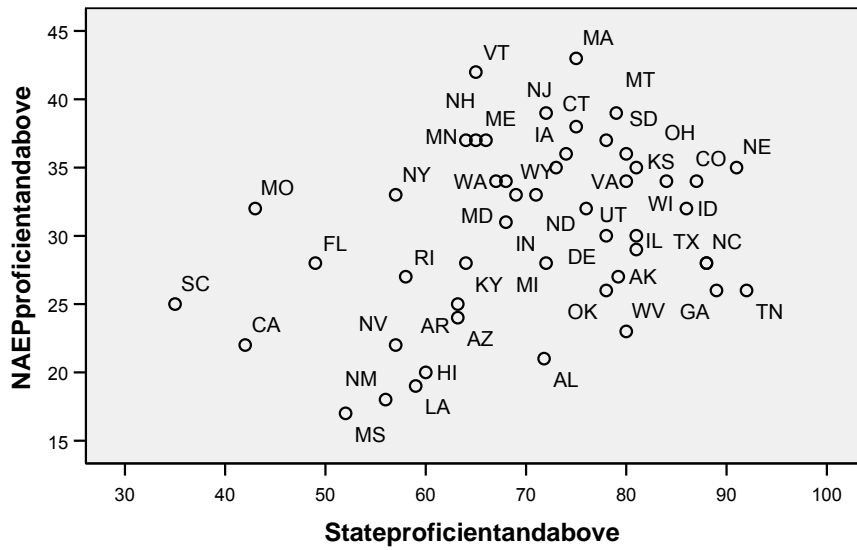


Figure 2. Scatter Plot of State Proficient and Above and NAEP Proficient and Above

A correlation coefficient was computed to determine if a linear relationship existed between state proficient and above and NAEP proficient and above data. In the Model Summary (Table 1) the Pearson correlation coefficient ( $r=.327$ ,  $p\text{-value} = 0.021$ ), indicated a positive linear relationship between state proficient and above and NAEP proficient and above data.

Table 1

Model Summary of State and NAEP (Proficient and Above)

Model	R	R Square	P-value
1	.327(a)	.107	.021

a Predictors: (Constant State Proficient and Above)

The following linear model was constructed:  $\hat{y} = 19.272 + .157x$ , where  $x$  = state proficient and above and  $y$  = NAEP proficient and above. A regression analysis was conducted which yielded a significance level of .021 for the independent variable (state proficient and above). This finding indicated that the linear model was statistically significant in predicting the state proficient and above scores from the NAEP proficient and above scores, as displayed in Table 2. Using State proficient and above scores, 10% of the variability in NAEP proficient and above scores were predicted by the regression model.

Table 2

Regression Model of State and NAEP (Proficient and Above)

Mode 1		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	19.272	4.717		4.086	.000
	State Proficient and Above	.157	.066	.327	2.395	.021

a Dependent Variable: NAEP Proficient and Above

### Research Question 2

What comparison can be made, if any, between the percent of eighth grade students identified as proficient and above in reading on state assessments and on the NAEP in the four census-defined regions identified by the NAEP in 2007?

The desired outcome was to gain insight as to whether eighth grade students' performance data from 2007, disaggregated by census-defined regions on state reading

exams and on the NAEP reading exam, could be compared. Data were collected from the 2006-2007 CSPR and from the NAEP State Snapshot Reports on each individual state and compiled into an SPSS spreadsheet, as shown in Appendix F.

Initially, the researcher sought to compare the mean proficiency level of NAEP and the state data by region; however, when histograms of each region were developed it was deemed that the data did not have a normal distribution, as demonstrated in the histograms located in Appendixes I- L.

All four regions revealed there was a difference in the probability distribution between state proficient and above and NAEP proficient and above data. A Wilcoxon nonparametric test was conducted since the aforementioned data were not considered normal, as demonstrated in Tables 3-6.

Table 3

Wilcoxon (Midwest)

	State Proficient and Above – NAEP Proficient and Above
Z	-3.064(a)
Asymp. Sig. (2-tailed)	.002

a Based on negative ranks.

b Wilcoxon Signed Ranks Test



Table 4

Wilcoxon (Northeast)

	State Proficient and Above – NAEP Proficient and Above
Z Asymp. Sig. (2-tailed)	-2.666(a) .008

a Based on negative ranks.

b Wilcoxon Signed Ranks Test

Table 5

Wilcoxon (South)

	State Proficient and Above – NAEP Proficient and Above
Z Asymp. Sig. (2-tailed)	-3.517(a) .000

a Based on negative ranks.

b Wilcoxon Signed Ranks Test

Table 6

Wilcoxon (West)

	State Proficient and Above – NAEP Proficient and Above
Z Asymp. Sig. (2-tailed)	-3.182(a) .001

a Based on negative ranks.

b Wilcoxon Signed Ranks Test

The data outputs revealed the ranked scores for the 2007 State Proficient and Above data and NAEP Proficient and Above data are different in each region (p-values < 0.05).

### Research Question 3

What relationship exists, if any, between the percent of low socioeconomic eighth grade students identified as proficient and above in reading on state assessments and on the NAEP, in the four census-defined regions identified by the NAEP in 2007?

The desired outcome was to gain insight as to whether a relationship existed between low socioeconomic eighth grade students' performance on state reading exams and on the NAEP reading exam, by census-defined regions, in 2007. Data were collected from the 2006-2007 CSPR and from the NAEP State Snapshot Report on each individual state and compiled into an SPSS spreadsheet, as shown in Appendix G.

#### Low Socioeconomic Census-Defined Region (Midwest)

The researcher initially established a scatter plot, to determine whether a linear relationship was present between low socioeconomic state proficient and above and low socioeconomic NAEP proficient and above data in the Midwest census-defined region. Based on the scatter plot, the low socioeconomic state proficient and above and low socioeconomic NAEP proficient and above data in the Midwest census-defined region appeared to have a weak positive linear relationship as displayed in Figure 3.

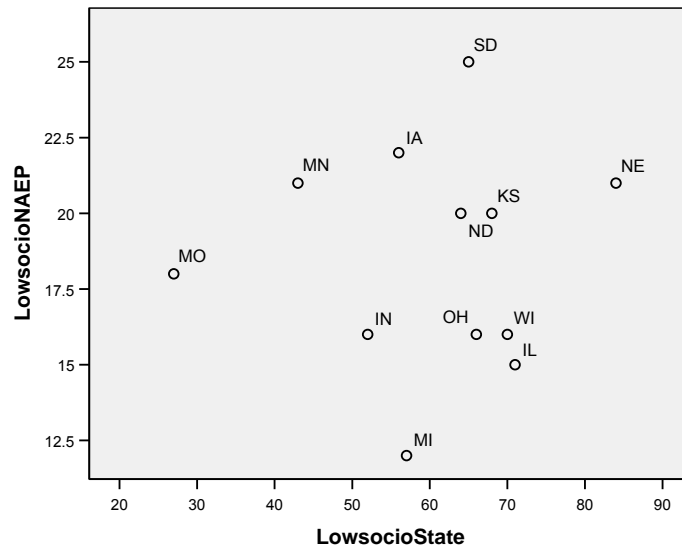


Figure 3. Scatter Plot of Midwest Census-Defined Region (Low Socioeconomic)

A correlation coefficient was conducted to determine if a linear relationship existed between low socioeconomic state proficient and above and low socioeconomic NAEP proficient and above data in the Midwest census-defined region. In the Model Summary (Table 7) the Pearson correlation coefficient was ( $r = 0.053$ ,  $p\text{-value} = 0.869$ ), which indicated no statistically significant linear relationship between low socioeconomic state proficient and above and low socioeconomic NAEP proficient and above data in the Midwest census-defined region.

Table 7

Model Summary of Low Socioeconomic (Midwest)

Model	R	R Square	P-value
1	.053(a)	.003	.869

a Predictors: (Constant), Low Socio State

No linear model was constructed because this finding indicated that the linear model was not statistically significant in predicting the low socioeconomic state proficient and above from the low socioeconomic NAEP proficient and above data in the Midwest census-defined region.

Low Socioeconomic Census-Defined Region (Northeast)

The researcher initially established a scatter plot, which was useful to determine if a linear relationship was present between low socioeconomic state proficient and above and low socioeconomic NAEP proficient and above data in the Northeast census-defined region. Based on the scatter plot, the low socioeconomic state proficient and above and low socioeconomic NAEP proficient and above data in the Northeast census-defined region appeared to have a weak positive linear relationship, as displayed in Figure 4.

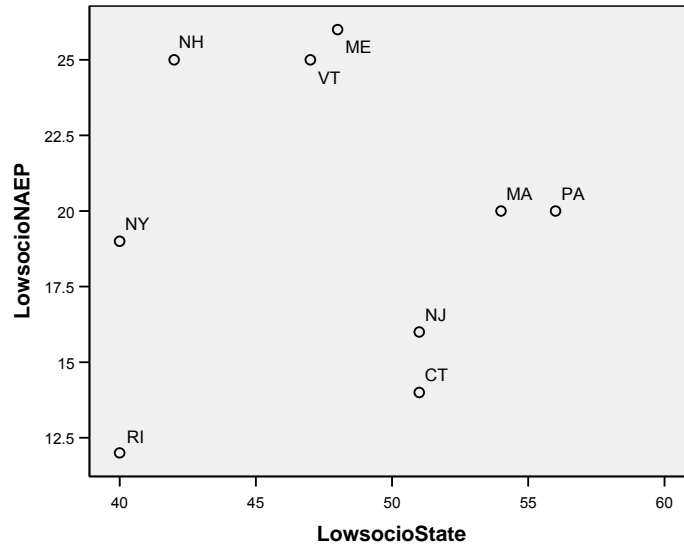


Figure 4. Scatter Plot of Northeast Census-Defined Region (Low Socioeconomic)

A correlation coefficient was computed to determine if a linear relationship existed between low socioeconomic state proficient and above and low socioeconomic NAEP proficient and above data in the Northeast census-defined region. In the Model Summary (Table 8) the Pearson correlation coefficient was ( $r = 0.025$ ,  $p\text{-value} = .949$ ), which indicated no statistically significant linear relationship between low socioeconomic state proficient and above and low socioeconomic NAEP proficient and above data in the Northeast census-defined region.

Table 8

Model Summary of Low Socioeconomic (Northeast)

Model	R	R Square	P-value
1	.025(a)	.001	.949

a Predictors: (Constant), Low Socioeconomic State

No linear model was constructed because this finding indicated that the linear model was not statistically significant in predicting the low socioeconomic state proficient and above from the low socioeconomic NAEP proficient and above data in the Northeast census-defined region.

Low Socioeconomic Census-Defined Region (South)

The researcher initially established a scatter plot, which was significant in determining whether a linear relationship was present between low socioeconomic state proficient and above and low socioeconomic NAEP proficient and above data in the South census-defined region. Based on the scatter plot, the low socioeconomic state proficient and above and low socioeconomic NAEP proficient and above data in the South census-defined region appeared to have a weak positive linear relationship, as displayed in Figure 5.

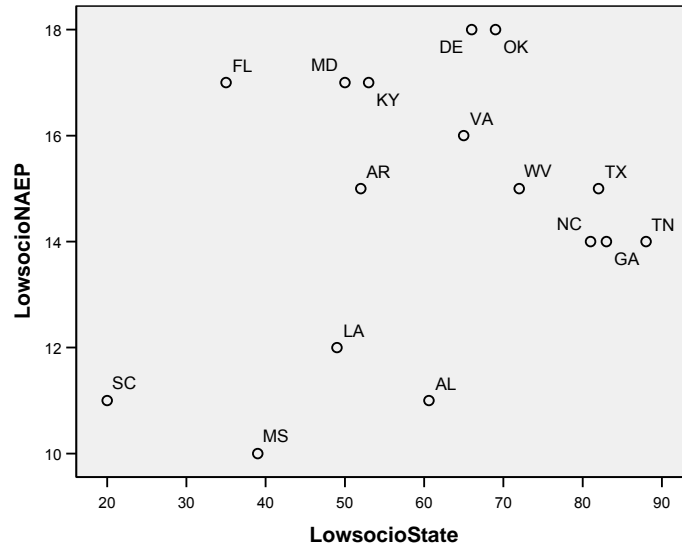


Figure 5. Scatter Plot of South Census-Defined Region (Low Socioeconomic)

A correlation coefficient was then computed to determine if a linear relationship existed between low socioeconomic state proficient and above and low socioeconomic NAEP proficient and above data in the South census-defined region. In the Model Summary (Table 9) the Pearson correlation coefficient was ( $r = 0.260$ ,  $p\text{-value} = 0.330$ ), which indicated no statistically significant linear relationship between low socioeconomic state proficient and above and low socioeconomic NAEP proficient and above data in the South census-defined region.

Table 9

Model Summary of Low Socioeconomic (South)

Model	R	R Square	P-value
1	.260(a)	.068	.330

a Predictors: (Constant), Low Socioeconomic State

No linear equation was constructed because this finding indicated that a linear model would not be statistically significant in predicting the low socioeconomic state proficient and above from the low socioeconomic NAEP proficient and above in the South census-defined region.

Low Socioeconomic Census-Defined Region (West)

The researcher initially established a scatter plot, which was useful in determining whether a linear relationship was present between low socioeconomic state proficient and above and low socioeconomic NAEP proficient and above data in the West census-defined region. Based on the scatter plot, the low socioeconomic state proficient and above scores and low socioeconomic NAEP proficient and above scores in the West census-defined region, appeared to have a moderately strong positive linear relationship as displayed in Figure 6.



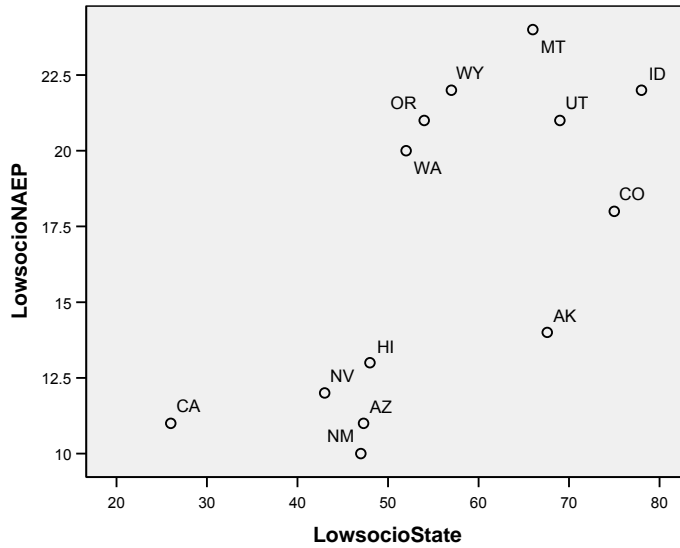


Figure 6. Scatter Plot of West Census-Defined Region (Low Socioeconomic)

A correlation coefficient was computed to determine if a linear relationship existed between low socioeconomic state proficient and above and low socioeconomic NAEP proficient and above data in the West census-defined region. In the Model Summary (Table 10) the Pearson correlation coefficient ( $r = .665$ ,  $p\text{-value} = 0.013$ ), indicated a statistically significant linear relationship between low socioeconomic state proficient and above and low- socioeconomic NAEP proficient and above data in the West census-defined region

Table 10

Model Summary of Low Socioeconomic (West)

Model	R	R Square	P-value
1	.665(a)	.442	.013

a Predictors: (Constant), Low Socioeconomic State

The following linear model was constructed:  $\hat{y} = 3.820 + .232x$ , where  $x$  = low socioeconomic proficient and above state in the West census-defined region and  $y$  = low socioeconomic proficient and above NAEP in the West census-defined region. A regression analysis was conducted, which yielded a significance level of .013. This finding indicated that the linear model was statistically significant in predicting the low socioeconomic NAEP proficient and above from the low socioeconomic state proficient and above data in the West census-defined region, as displayed in Table 11. Using low socioeconomic state proficient and above test scores, 44.2% of the variability in low socioeconomic NAEP proficient and above scores was determined by the regression model.

Table 11

Regression Model of Low Socioeconomic (West)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	3.820	4.549		.840	.419
	Low Socio State	.232	.079	.665	2.951	.013

a Dependent Variable: Low Socioeconomic NAEP

#### Research Question 4

What relationship exists, if any, between the percent of nonwhite eighth grade students identified as proficient and above in reading on state assessments and on the NAEP, in the four census-defined regions identified by NAEP in 2007?

The desired outcome was to gain insight as to whether there was a relationship between nonwhite eighth grade students' performance on state reading exams and on the NAEP reading exam, by census-defined regions, in 2007. Data were collected from the 2006-2007 CSPR and from the NAEP State Snapshot Report on each individual state and compiled into an SPSS spreadsheet, as shown in Appendix H.

#### Nonwhite Census-Defined Region (Midwest)

The researcher initially established a scatter plot, which was useful in determining whether a linear relationship was present between nonwhite state proficient and above and nonwhite NAEP proficient and above data in the Midwest census-defined region. Based on the scatter plot, the nonwhite state proficient and above and nonwhite NAEP proficient and above data in the Midwest census-defined region appeared to have a weak

positive linear relationship, as displayed in Figure 7. Illinois was determined to be an influential point due to it being an outlier.

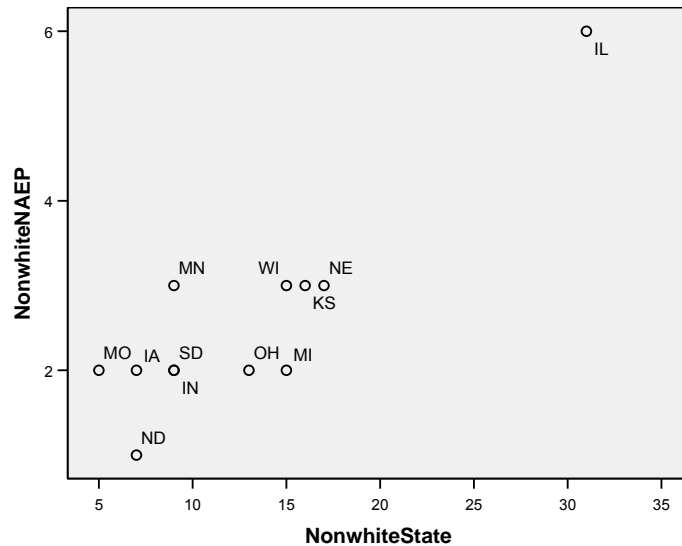


Figure 7. Scatter Plot of Midwest Census-Defined Region (Nonwhite)

A correlation coefficient was computed to determine if a linear relationship existed between nonwhite state proficient and above and nonwhite NAEP proficient and above data in the Midwest census-defined region. In the Model Summary (Table 12) the Pearson correlation coefficient ( $r = .887$ ,  $p\text{-value} < 0.01$ ), indicated a statistically significant linear relationship between nonwhite state proficient and above and nonwhite NAEP proficient and above data in the Midwest census-defined region.

Table 12

Model Summary of Nonwhite (Midwest)

Model	R	R Square	P-value
1	.887(a)	.786	<.01

a Predictors: (Constant), Nonwhite State

The following linear model was constructed:  $\hat{y} = .583 + .157x$ , where  $x =$  nonwhite proficient and above state in the Midwest census-defined region and  $y =$  nonwhite proficient and above NAEP in the Midwest census-defined region. A regression analysis was conducted, which yielded a significance level of approximately .000. This finding indicated that the linear model was statistically significant in predicting the nonwhite state proficient and above from the nonwhite NAEP proficient and above data in the Midwest census-defined region, as displayed in Table 13. Using state proficient and above scores, 78.6% of the variability in NAEP proficient and above scores was determined by the regression model.

Table 13

Regression Model of Nonwhite (Midwest)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.583	.373		1.563	.149
	Nonwhite State	.157	.026	.887	6.059	.000

a Dependent Variable: Nonwhite NAEP

Nonwhite Census-Defined Region (Northeast)

The researcher initially established a scatter plot, which was useful in determining whether a linear relationship was present between nonwhite state proficient and above and nonwhite NAEP proficient and above data in the Northeast census-defined region. Based on the scatter plot, the nonwhite state proficient and above and nonwhite NAEP proficient and above data in the Northeast census-defined region appeared to have a strong positive linear relationship, as displayed in Figure 8.

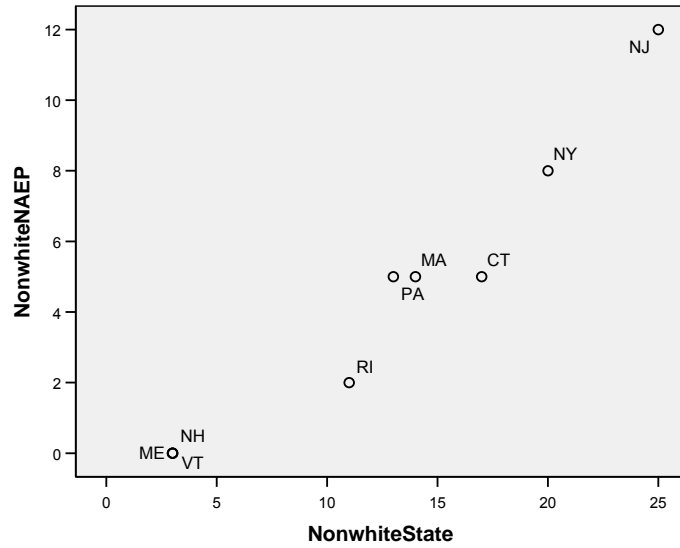


Figure 8. Scatter Plot of Northeast Census-Defined Region (Nonwhite)

A correlation coefficient was computed to determine if a linear relationship existed between nonwhite state proficient and above and nonwhite NAEP proficient and above data in the Northeast census-defined region. In the Model Summary (Table 14) the Pearson correlation coefficient was ( $r = .971$ ,  $p\text{-value} < 0.01$ ), which indicated a statistically significant positive linear relationship between nonwhite state proficient and above and nonwhite NAEP proficient and above data in the Northeast census-defined region.

Table 14

Model Summary of Nonwhite (Northeast)

Model	R	R Square	P-value
1	.971(a)	.943	<.01

a Predictors: (Constant), Nonwhite State

The following linear model was constructed:  $\hat{y} = -1.955 + .501x$ , where  $x =$  nonwhite proficient and above state in the Northeast census-defined region and  $y =$  nonwhite proficient and above NAEP in the Northeast census-defined region. A regression analysis was conducted, which yielded a significance level of approximately .000. This finding indicated that the linear model was statistically significant in predicting the nonwhite NAEP proficient and above from the nonwhite state proficient and above data in the Northeast census-defined region, as displayed in Table 15. Using nonwhite state proficient and above scores, 94.3%, of the variability in nonwhite NAEP proficient and above scores were determined by the regression model.



Table 15

Regression Model of Nonwhite (Midwest)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-1.955	.665		-2.941	.022
	Nonwhite State	.501	.047	.971	10.737	.000

a Dependent Variable: Nonwhite NAEP

Nonwhite Census-Defined Region (South)

The researcher initially established a scatter plot, which was useful in determining whether a linear relationship was present between nonwhite state proficient and above and nonwhite NAEP proficient and above data in the South census-defined region. Based on the scatter plot, the nonwhite state proficient and above and nonwhite NAEP proficient and above data in the South census-defined region appeared to have a moderately strong positive linear relationship, as displayed in Figure 9.

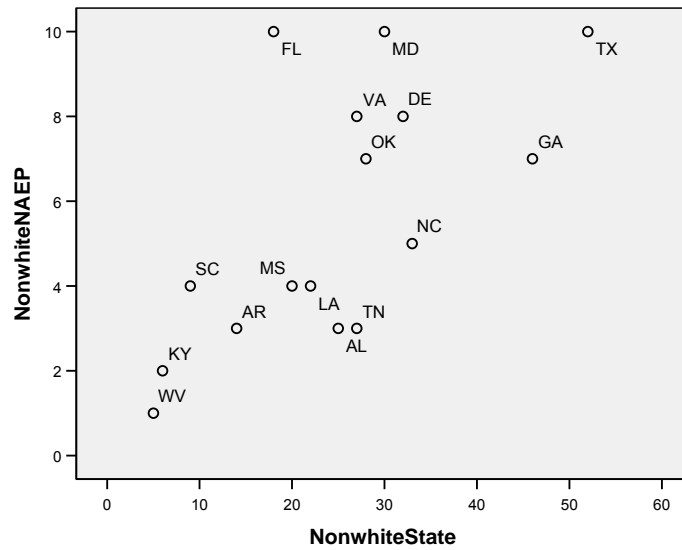


Figure 9. Scatter Plot of South Census-Defined Region (Nonwhite)

A correlation coefficient was then computed to determine if a linear relationship existed between nonwhite state proficient and above and nonwhite NAEP proficient and above data in the South census-defined region. In the Model Summary (Table 16) the Pearson correlation coefficient was ( $r = 0.650$ ,  $p\text{-value} = .006$ ), which indicated a statistically significant positive linear relationship between nonwhite state proficient and above and nonwhite NAEP proficient and above data in the South census-defined region.

Table 16

Model Summary of Nonwhite (South)

Model	R	R Square	P-value
1	.650(a)	.423	.006

a Predictors: (Constant), Nonwhite State

The following linear model was constructed:  $\hat{y} = 1.861 + .150x$ , where  $x$  = nonwhite proficient and above state in the South census-defined region and  $y$  = nonwhite proficient and above NAEP in the South census-defined region. A regression analysis was conducted, which yielded a significance level of .006. This finding indicated that the linear model was statistically significant in predicting the nonwhite NAEP proficient and above from the South state proficient and above data in the South census-defined Region, as displayed in Table 17. Using nonwhite state proficient and above scores, 42.3% of the variability in nonwhite NAEP proficient and above scores was determined by the regression model.

Table 17

Regression Model (South)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	1.861	1.299		1.433	.174
	Nonwhite State	.150	.047	.650	3.203	.006

a Dependent Variable: Non White NAEP

### Nonwhite Census-Defined Region (West)

The researcher initially established a scatter plot, which was useful in determining whether a linear relationship was present between nonwhite state proficient and above and nonwhite NAEP proficient and above data in the West census-defined region. Based on the scatter plot, the nonwhite state proficient and above and nonwhite NAEP proficient and above data in the West census-defined region appeared to have a positive linear relationship, as displayed in Figure 10.

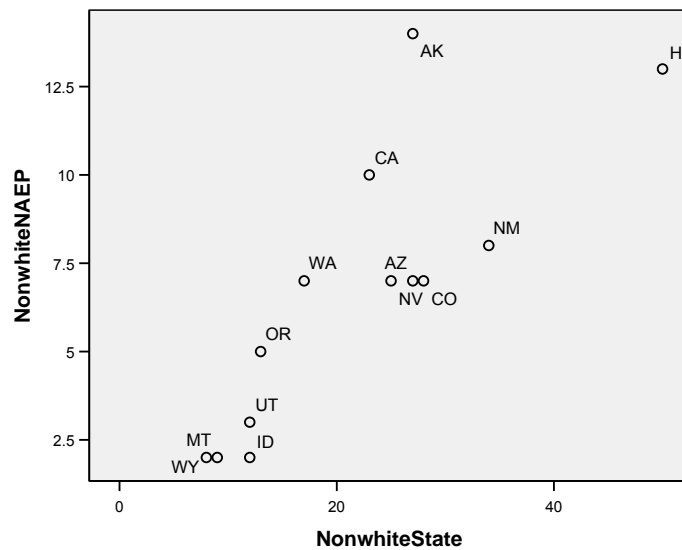


Figure 10. Scatter Plot of West Census-Defined Region (Nonwhite)

A correlation coefficient was computed to determine if a linear relationship existed between nonwhite state proficient and above and nonwhite NAEP proficient and above data in the West census-defined region. In the Model Summary (Table 18) the Pearson correlation coefficient was ( $r = 0.814$ ,  $p\text{-value} = 0.001$ ), which indicated a

statistically significant positive linear relationship between nonwhite state proficient and above and nonwhite NAEP proficient and above data in the West census-defined region.

Table 18

Model Summary of Nonwhite (West)

Model	R	R Square	P-Value
1	.814(a)	.663	.001

a Predictors: (Constant), Nonwhite State

The following linear model was constructed:  $\hat{y} = .733 + .272x$ , where  $x =$  nonwhite proficient and above state in the West census-defined Region and  $y =$  nonwhite proficient and above NAEP in the West census-defined region. A regression analysis was conducted, which yielded a significance level of .001. This finding indicated that the linear model was statistically significant in predicting the nonwhite NAEP proficient and above from the nonwhite state proficient and above data in the West census-defined region, as displayed in Table 19. Using nonwhite state proficient and above scores, 66.3% of the variability in nonwhite NAEP proficient and above scores was determined by the model.

Table 19

Regression Model of Nonwhite (West)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.733	1.443		.508	.621
	NonwhiteState	.272	.058	.814	4.656	.001

a Dependent Variable: Nonwhite NAEP

### Summary

The four research questions provided the framework for the analysis of the data presented in Chapter 4. A discussion of the results, as well as conclusions, implications for practice, and recommendations for future research follow in the next chapter.

## CHAPTER FIVE: CONCLUSIONS AND RECOMMENDATIONS

### Introduction

This investigation was conducted to offer insight into the testing provisions of the No Child Left Behind (NCLB). According to Shaul and Ganson (2005), NCLB was a Congressional effort to bring greater accountability and improvement in student achievement to the forefront in the educational reform movement (Shaul & Ganson). An increased transparency of student performance data and the lack of uniform accountability systems created difficulty for educators, policy-makers and the general public to make state-by-state comparisons in the area of reading (National Center for Educational Statistics, 2007).

To keep the enormity of the NCLB's policies manageable, two titles were chosen to guide the analysis of data for the research: Title I – Improving the Academic Achievement of the Disadvantaged and Title VI – Flexibility and Accountability. The main purpose of Title I was to “ensure that all children have a fair, equal, and significant opportunity to obtain a high quality education and reach, at minimum, proficiency on challenging state academic standards and state assessments” (United States Department of Education, 2005). The principal intent of Title VI was to create flexible environments in which states, Local Education Agencies (LEAs) and higher institutions developed assessments to elevate student performance and close the achievement gap (United States Department of Education, 2007b).

The aim of this study was to provide information for policymakers and practitioners by comparing the performance of eighth grader students in 2007 on state standardized reading assessments and in the four census-defined regions identified by National Assessment of Educational Progress (NAEP). The review of literature presented a context for this study. In its 2002 report, the Ad Hoc Committee on Confirming Test Results stated that “differences between the National Assessment of Educational Progress and state testing programs must be explored and reported” (Stoneberg, 2007), which clearly warranted further investigation.

In 2007, the National Center for Educational Statistics (NCES), a division of the United States Department of Education (USDOE), administered and scored the NAEP exam in reading by randomly selecting eighth grade students in all 50 states to achieve a representative population of the country (United States Department of Education, 2008b). The data were reported as part of the Nation’s Report Card.

Sec. 1111(b) (l) (A) of the NCLB required states to set demanding content standards as well as challenging student achievement standards (Linn, 2003). States were required to establish measurement tools to evaluate whether LEAs were making progress with all students (United States Department of Education, 2004d). In 2007, the 50 states of the United States conducted eighth grade reading assessments and reported data related to the tests.

Chapter 5 provides the results and conclusions of the study and discusses how the data presented in Chapter 4 relate to each of the four research questions.



Recommendations for future research and concluding comments are presented by the researcher.

### Statement of the Problem

The purpose of this study was to compare the performance of eighth grade students in 2007 on state standardized reading tests and in the four census-defined regions identified by NAEP. The researcher chose to compare all states due to an NCLB requirement that all states report eighth grade student reading proficiency data and the availability of NAEP scores of eighth grade students by state. Furthermore, this study attempted to show that a national assessment system could provide more accurate comparisons between states as a means of validating the provisions of NCLB.

### Research Questions

The study was guided by the following research questions:

1. What relationship exists, if any, between the percent of eighth grade students identified as proficient in reading on state assessments and on the NAEP in 2007?
2. What comparison can be made, if any, between the percent of eighth grade students identified as proficient and above in reading on state assessments and on the NAEP, in the four census-defined regions identified by the NAEP in 2007?
3. What relationship exists, if any, between the percent of low socioeconomic eighth grade students identified as proficient and above in reading on state assessments and on the NAEP, in the four census-defined regions identified by the NAEP in 2007?
4. What relationship exists, if any, between the percent of nonwhite eighth grade students identified as proficient and above in reading on state

assessments and on the NAEP, in the four census-defined regions identified by NAEP in 2007?

### Data Collection

In November 2008, eighth grade reading performance data were collected from the NCES 2007 State Snapshot Report (Appendix A) and from SY 2006-2007 Consolidated State Performance Reports (Appendix B). The data were organized in an SPSS worksheet and disaggregated by the four census-defined regions identified by NAEP (Figure 1). Data were further disaggregated by low socioeconomic students and nonwhite students. This information was analyzed using several statistical procedures to answer the four research questions.

### Instrument Reliability and Validity

#### Background

The federal government contracted with the Education Commission on States and created the NAEP in 1969. The goal of the establishment of NAEP was to monitor achievement in 10 learning areas and to assess change in achievement over the years. Information gathered from NAEP was to then be reviewed and used by policymakers to institute change in public education (Resnick, 1980).

The accountability provisions of Title I were strengthened under NCLB to provide a mechanism for holding states, school districts, and schools accountable for improving the academic achievement of all students (Institute of Educational Sciences, 2007). A major provision of NCLB was the expectation that states set demanding content

standards as well as challenging student achievement standards (Sec. 11 1 l(b)(1)(A), as cited in (Linn, 2003). Determining Adequate Yearly Progress (AYP) was the means by which a state measures its demanding content and challenging student achievement standards. Each state establishes its own criteria when assessing a student's yearly progress and must report these scores to parents. The federal government requires states to have 100% of their students demonstrate success on state-developed standardized tests by the year 2014 (United States Department of Education, 2004d).

### Reliability

According to Jones and Olkin (2004), the NAEP was considered the “Gold Standard” in measuring student achievement. The Education Testing Service (ETS) was responsible for the design, analysis and reporting of the 2007 NAEP reading data. NAEP assessments contain both multiple-choice and constructed response questions. Constructed response questions require students to provide their own answers. In order to score large numbers of constructed response questions, over 3 million annually, with a high level of accuracy, NAEP incorporated extensive quality control measures to ensure reliability (United States Department of Education, 2008a). It was incumbent on each state to develop its own procedures to ensure accurate and reliable design, analysis and reporting of state assessments under the terms of NCLB.

## Summary and Discussion

The following section provides analysis and discussion of each of the four research questions.

### Research Question 1

What relationship exists, if any, between the percentage of eighth grade students identified as proficient in reading on state assessments and on the NAEP in 2007?

The desired outcome was to gain insight as to whether there was a relationship between eighth grade student performance on state reading assessments in 2007 and eighth grade student reading performance on the NAEP in 2007. To analyze Research Question 1, eighth grade reading performance data were gathered from two sources. State data were amassed from the United States Department of Education's 2006-2007 CSPR. NAEP data were collected from the National Center for Education Statistics (NCES). A scatter plot was computed to determine if there was a relationship between eighth grade students identified as proficient and above in reading on state assessments and on the NAEP assessment in 2007. The scatter plot showed performance on state assessments would not necessarily predict performance on the NAEP assessment, a factor which prompted further investigation. A Pearson correlation coefficient ( $r = .327$ ,  $p\text{-value} = 0.021$ ) indicated statistical significance that a positive linear relationship was present between state proficient and above and NAEP proficient and above scores. The researcher established a linear model  $\hat{y} = 19.272 + .157x$ , where  $x$  = state proficient and above and  $y$  = NAEP proficient and above. The linear model indicated there was a way to

predict student performance on the NAEP by analyzing student performance on state assessments.

### Discussion

Under the provisions of NCLB, states were given the autonomy to set standards and to develop standardized tests to measure student performance. When analyzing the data, all states reported higher percentages of students performing at the proficient and above levels than on the NAEP. Fourteen states reported more than a 50 percentage point difference in eighth grade reading scores on state standardized tests than on the NAEP. Six states reported less than a 25 percentage point difference in state eighth grade test scores than on the NAEP. The difference in percentages could be attributed, but not limited to a variety of factors such as dissimilarity in standards being taught, populations served, or the ease of the state standardized test.

In analyzing data on a state level, Tennessee reported the greatest difference in the percentage of students performing at proficient and above on the state's standardized test at 92% in contrast to 26% of its students scoring at proficient and above on the NAEP. This represented a 66 percentage point difference. South Carolina reported the least amount of difference between performance on state exams versus the NAEP with 35% of its students scoring at proficient and above on its state exam and 25% on the NAEP. This represented a 10 percentage point difference. Percentage point differences between states could be attributed, but not limited to factors to the aforementioned. As stated, the established linear model allowed stakeholders to predict performance on the

NAEP by using state test scores which could promote differences in performance to be analyzed and minimized.

## Research Question 2

What comparison can be made, if any, between the percentage of eighth grade students identified as proficient in reading on state assessments and on the NAEP, in the four census-defined regions identified by the NAEP in 2007?

The desired outcome was to gain insight as to whether eighth grade students' performance data from 2007, disaggregated by census-defined regions on state reading exams and on the NAEP reading exam, could be compared. Data were collected from the 2006-2007 CSPR and from the NAEP State Snapshot Report on each individual state and compiled into an SPSS.

Initially, the researcher sought to compare the mean proficiency level of NAEP and the state data by region; however, when histograms of each region were created, it was deemed that the data did not have a normal distribution. A Wilcoxon test was conducted, which concluded the ranked scores for the 2007 state proficient and above data and for the NAEP proficient and above data were different in each region (p-values<.05).

## Discussion

The researcher analyzed the data by regions to see if any insight could be gained on how regional data could guide stakeholders in making comparisons between eighth grade students performance on state reading exams and on the NAEP. Due to the results

of the data not being normal and the findings of the Wilcoxon tests, averages of the regional scores were compared (Appendix M). The state proficient and above average was 74% in the thirteen states that comprised the Midwest region while the NAEP proficient and above average for the region was 33%. The nine states in the Northeast region reported a 67% average for state proficient and above and a 37% average for NAEP proficient and above. In the sixteen states in the South region, an average of 71% of the students scored at state proficient and above and 26% for NAEP. Thirteen states from the West region reported an average of 69% of the students scored at proficient and above on state tests and an average of 28% was reported for the NAEP.

Additional regional data were compared to garner information (Appendix M). The state proficient and above average for all regions was 71% while the NAEP proficient and above average was 34%. The Northeast region reported the closest average difference between state proficient and above scores at 31 % while the South region had the greatest average difference at 45%. Closer analysis showed students from all regions demonstrated higher state proficient and above scores than NAEP proficient and above scores. Populations served, state standards, staff development and pedagogical strategies could be, but are not limited to, important factors that impacted the findings.

### Research Question 3

What relationship exists, if any, between the percentage of low socioeconomic eighth grade students identified as proficient and above in reading on state assessments and on the NAEP, in the four census-defined regions identified by the NAEP in 2007?

The desired outcome was to gain insight as to whether there was a relationship between low socioeconomic eighth grade students' performance on state reading exams and on the NAEP reading exam, by census-defined regions, in 2007. Data were collected from the 2006-2007 CSPR and from the NAEP State Snapshot Report on each individual state and compiled into an SPSS spreadsheet. Each region was analyzed separately by using a similar set of statistical tests for consistency.

#### Low Socioeconomic Census-Defined Region (Midwest)

A scatter plot was computed to determine if there was a relationship between low socioeconomic eighth grade students from the Midwest census-defined region identified as proficient and above in reading on state assessments and on the NAEP assessment in 2007. The scatter plot showed low socioeconomic state proficient and above and low socioeconomic NAEP proficient and above data in the Midwest census-defined region appeared to have a weak positive linear relationship. As a result, performance on state assessments would not necessarily predict performance on the NAEP assessment, a factor which prompted further investigation. A Pearson correlation coefficient ( $r = 0.053$ ,  $p\text{-value} = 0.869$ ) indicated that no statistically significant linear relationship was present between low socioeconomic state proficient and above and low socioeconomic NAEP proficient and above data in the Midwest census-defined region.



#### Low Socioeconomic Census-Defined Region (Northeast)

A scatter plot was computed to determine if there was a relationship between low socioeconomic economic eighth grade students from the Northeast census-defined region identified as proficient and above in reading on state assessments and on the NAEP assessment in 2007. The scatter plot showed low socioeconomic state proficient and above and low socioeconomic NAEP proficient and above data in the Northeast census-defined region appeared to have a weak positive linear relationship. As a result, performance on state assessments would not necessarily predict performance on the NAEP assessment, a factor which prompted further investigation. A Pearson correlation coefficient ( $r = 0.025$ ,  $p\text{-value} = 0.949$ ) indicated that no statistically significant linear relationship was present between Low socioeconomic state proficient and above and low socioeconomic NAEP proficient and above data in the Midwest census-defined region.

#### Low Socioeconomic Census-Defined Region (South)

A scatter plot was computed to determine if there was a relationship between low socioeconomic economic eighth grade students from the South census-defined region identified as proficient and above in reading on state assessments and on the NAEP assessment in 2007. The scatter plot showed low socioeconomic state proficient and above and low socioeconomic NAEP proficient and above data in the South census-defined region appeared to have a weak positive linear relationship. As a result, performance on state assessments would not necessarily predict performance on the NAEP assessment, a factor which prompted further investigation. A Pearson correlation

coefficient ( $r = 0.260$ ,  $p\text{-value} = 0.330$ ) indicated that no statistically significant linear relationship was present between Low socioeconomic state proficient and above and low socioeconomic NAEP proficient and above data in the Midwest census-defined region.

#### Low Socioeconomic Census-Defined Region (West)

A scatter plot was computed to determine if there was a relationship between low socioeconomic eighth grade students from the West census-defined region identified as proficient and above in reading on state assessments and on the NAEP assessment in 2007. The scatter plot showed low socioeconomic state proficient and above and low socioeconomic NAEP proficient and above data in the West census-defined region appeared to have a moderately strong positive linear relationship. As a result, performance on state assessments could predict performance on the NAEP assessment, a factor which prompted further investigation. A Pearson correlation coefficient ( $r = .665$ ,  $p\text{-value} = 0.013$ ) indicated that a statistically significant linear relationship was present between low socioeconomic state proficient and above and low socioeconomic NAEP proficient and above data in the West census-defined region. The researcher established a linear model  $\hat{y} = 3.820 + .232x$ , where  $x$  = low socioeconomic state proficient and above in the West census-defined region and  $y$  = low socioeconomic NAEP proficient and above in the West census-defined region. The linear model indicated there was a way to predict low socioeconomic economic student performance on the NAEP by analyzing low socioeconomic economic student performance on state assessments.

## Discussion

The statistical procedures determined that the West census-defined region low socioeconomic state reading scores were able to predict performance on the NAEP; however, the Midwest, Northeast and South low socioeconomic census-defined regional scores were unable to predict performance on the NAEP. Factors such as populations served, use of Title I funding, professional develop and pedagogy could have contributed to the reason the West census-defined region was able to predict performance while the other regions were not.

Analysis of the regional data showed the average score reported by states in the Midwest region for low socioeconomic state proficient and above was 60% while the low socioeconomic NAEP proficient and above was 19%. Missouri's low socioeconomic students had the least amount of difference between state exams and the NAEP at 9% while the greatest difference was Nebraska's 63%. In the Northeast region the average score reported by states for their low socioeconomic students on state exams was 48% while 20% was the reported average for low socioeconomic students on the NAEP. New Hampshire's low socioeconomic students had the least number of difference between state exams and the NAEP at 17% while the greatest difference was Connecticut's 37%. In the South region the average score reported for low socioeconomic students was 60% on state tests and 15% on the NAEP. South Carolina's low socioeconomic students had the smallest difference between state exams and the NAEP at 9% while the greatest difference was Tennessee's 74% (Appendix M).

The West region reported an average score for low socioeconomic students on state exams at 56% on the NAEP at 17%. California's low socioeconomic students had the least amount of difference between state exams and the NAEP at 15% while the greatest difference was Colorado at 57% (Appendix M).

Further analysis of regional information showed that low socioeconomic students on average scored lower on state exams and on the NAEP when compared to the average performance of all students on the same exams. Seventy-one percent of all students scored at proficient and above on state exams and 34% scored at proficient or above on the NAEP. Fifty-seven percent of low socioeconomic students scored at proficient or above on state exams while 17% scored at proficient and above on the NAEP. Low socioeconomic students scored on average higher on state exams and on the NAEP than nonwhite students. On average 18% of nonwhite students scored at proficient and above on state tests while 5% of nonwhite students scored at proficient and above on the NAEP (Appendix M). How states set standards, establish assessments, use Title I funds, conduct staff development and establish pedagogy to affect low socioeconomic students could explain the findings.

#### Research Question 4

What relationship exists, if any, between the percentage of nonwhite eighth grade students identified as proficient and above in reading on state assessments and on the NAEP, in the four census-defined regions identified by the NAEP in 2007?

The desired outcome was to gain insight as to whether there was a relationship between nonwhite eighth grade students' performance on state reading exams and on the

NAEP reading exam, by census-defined regions, in 2007. Data were collected from the 2006-2007 CSPR and from the NAEP State Snapshot Report on each individual state and compiled into an SPSS spreadsheet. Each region was analyzed separately by using a similar set of statistical tests for consistency.

#### Nonwhite Census-Defined Region (Midwest)

A scatter plot was computed to determine if there was a relationship between nonwhite eighth grade students from the Midwest census-defined region identified as proficient and above in reading on state assessments and on the NAEP assessment in 2007. The scatter plot showed nonwhite state proficient and above and nonwhite NAEP proficient and above data in the West census-defined region appeared to have a moderately strong positive linear relationship. As a result, performance of nonwhite students on state assessments would not necessarily predict performance of nonwhite students on the NAEP assessment, a factor which prompted further investigation. A Pearson correlation coefficient ( $r = .887$ ,  $p\text{-value} = 0.01$ ) indicated that a statistically significant linear relationship was present between nonwhite state proficient and above and Nonwhite NAEP proficient and above data in the Midwest census-defined region. The researcher established a linear model  $\hat{y} = .583 + .157x$ , where  $x$  = nonwhite state proficient and above in the Midwest census-defined region and  $y$  = nonwhite NAEP proficient and above in the Midwest census-defined region. Using nonwhite state proficient and above scores, 78.6% of the variability in nonwhite NAEP proficient and above scores was determined by the regression model. The linear model indicated there

was a way to predict nonwhite student performance on the NAEP by analyzing nonwhite student performance on state assessments in the Midwest census-defined region.

#### Nonwhite Census-Defined Region (Northeast)

A scatter plot was computed to determine if there was a relationship between nonwhite eighth grade students from the Northeast census-defined region identified as proficient and above in reading on state assessments and on the NAEP assessment in 2007. The scatter plot showed nonwhite state proficient and above and nonwhite NAEP proficient and above data in the Northeast census-defined region appeared to have a moderately strong positive linear relationship. As a result, performance of nonwhite students on state assessments would possibly predict performance of nonwhite students on the NAEP assessment, a factor which prompted further investigation. A Pearson correlation coefficient ( $r = .971$ ,  $p\text{-value} = 0.01$ ) indicated that a statistically significant positive linear relationship was present between nonwhite state proficient and above and nonwhite NAEP proficient and above data in the Northeast census-defined region. The researcher established a linear model  $\hat{y} = 1.955 + .501x$ , where  $x$  = nonwhite state proficient and above in the Northeast census-defined region and  $y$  = nonwhite NAEP proficient and above in the Northeast census-defined region. Using nonwhite state proficient and above scores, 94.3%, of the variability in nonwhite NAEP proficient and above scores were determined by the regression model. The linear model indicated there was a way to predict nonwhite student performance on the NAEP by analyzing nonwhite student performance on state assessments, in the Northeast census-defined region.

### Nonwhite Census-Defined Region (South)

A scatter plot was computed to determine if there was a relationship between nonwhite eighth grade students from the South census-defined region identified as proficient and above in reading on state assessments and on the NAEP assessment in 2007. The scatter plot showed nonwhite state proficient and above and nonwhite NAEP proficient and above data in the South census-defined region appeared to have a moderately strong positive linear relationship. As a result, performance of nonwhite students on state assessments would not necessarily predict performance of nonwhite students on the NAEP assessment, a factor which prompted further investigation. A Pearson correlation coefficient ( $r = 0.650$ ,  $p\text{-value} = 0.006$ ) indicated that a statistically significant positive linear relationship was present between nonwhite state proficient and above and nonwhite NAEP proficient and above data in the Northeast census-defined region. The researcher established a linear model  $\hat{y} = 1.861 + .150x$ , where  $x$  = nonwhite state proficient and above in the South census-defined region and  $y$  = nonwhite NAEP proficient and above in the South census-defined region. Using nonwhite state proficient and above scores, 42.3%, of the variability in nonwhite NAEP proficient and above scores were determined by the regression model. The linear model showed there was a way to predict nonwhite student performance on the NAEP by analyzing nonwhite student performance on state assessments in the South census-defined region.

### Nonwhite Census-Defined Region (West)

A scatter plot was computed to determine if there was a relationship between nonwhite eighth grade students from the West census-defined region identified as proficient and above in reading on state assessments and on the NAEP assessment in 2007. The scatter plot showed nonwhite state proficient and above and nonwhite NAEP proficient and above data in the West census-defined region appeared to have a strong positive linear relationship. As a result, performance of nonwhite students on state assessments would not necessarily predict performance of nonwhite students on the NAEP assessment, a factor which prompted further investigation. A Pearson correlation coefficient ( $r = 0.814$ ,  $p\text{-value} = 0.001$ ) indicated that a statistically significant positive linear relationship was present between nonwhite state proficient and above and nonwhite NAEP proficient and above data in the Northeast census-defined region. The researcher established a linear model  $\hat{y} = .733 + .272x$ , where  $x$  = nonwhite state proficient and above in the West census-defined region and  $y$  = nonwhite NAEP proficient and above in the West census-defined region. Using nonwhite state proficient and above scores, 66.3%, of the variability in nonwhite NAEP proficient and above scores were determined by the regression model. The linear model showed there was a way to predict nonwhite student performance on the NAEP by analyzing nonwhite student performance on state assessments in the West census-defined region.



## Discussion

Analysis of the data showed nonwhite reading scores by state from all four census-defined regions could predict performance on the NAEP using the established linear model. By being able to predict NAEP scores, nonwhite student performance from each census-defined region can be compared at a national level.

Further analysis of each region illustrated the average score reported by states in the Midwest region for nonwhite state proficient and above was 11% while the nonwhite NAEP and above was 2%. Missouri's nonwhite students had the least amount of difference between state exams and the NAEP at 3% while the greatest difference was Nebraska's 25%. In the Northeast region the average score reported by states for their nonwhite students on state exams was 12% while 4% was the reported average for nonwhite students on the NAEP. Maine, New Hampshire and Vermont nonwhite students had the least amount of difference between state exams and the NAEP at 3% while the greatest difference was New Jersey's 12%. In the South region the average score reported for nonwhite students was 25% on state tests and 6% on the NAEP. Kentucky and West Virginia nonwhite students had the least amount of difference between state exams and the NAEP at 4% while the greatest difference was Texas' 42%. The West region reported an average score for nonwhite students on state exams at 22% on the NAEP at 7%. Wyoming's nonwhite socioeconomic students had the least amount of difference between state exams and the NAEP at 6% while the greatest difference was Hawaii at 37% (Appendix M).

Furthermore, data showed nonwhite students scored on average lower on state exams and on the NAEP when compared to all students in census-defined regions. Seventy-one percent of all students scored at proficient and above on state exams and 34% scored at proficient or above on the NAEP. Data disaggregated by regions showed 19% of nonwhite students scored at proficient and above on state tests while 5% of nonwhite students scored at proficient and above on the NAEP. Additionally, scores disaggregated by regions showed nonwhite students scored lower than low socioeconomic students. Fifty-seven percent of low socioeconomic students scored at proficient and above on state exams while 17% scored at proficient and above on NAEP exams (Appendix M). The above findings could have been influenced by a variety of, but not limited to, factors such as population's served, state standards and assessments, professional development and pedagogy.

### Conclusions

The literature review and research provided valuable insight into the testing provisions of NCLB. The statistical procedures conducted to analyze the data further expanded the study when it revealed the following:

1. A regression analysis using the linear model:

$\hat{y} = 19.272 + .157x$ , where  $x$  = state proficient and above and  $y$  = NAEP proficient and above, was statistically significant in predicting the state proficient and above scores from the NAEP proficient and above scores.

2. A Wilcoxon nonparametric test concluded that the ranked scores for state proficient and above data and NAEP proficient and above data were different in each region.
3. A regression analysis indicated no statistically significant linear relationship between low socioeconomic state proficient and above and low socioeconomic NAEP proficient and above data in the Midwest census-defined region.
4. A regression analysis indicated no statistically significant linear relationship between low socioeconomic state proficient and above and low socioeconomic NAEP proficient and above data in the Northeast census-defined region.
5. A regression analysis indicated no statistically significant linear relationship between low socioeconomic state proficient and above and low socioeconomic NAEP proficient and above data in the South census-defined region.
6. A regression analysis using the linear model:  
$$\hat{y} = 3.820 + .232x$$
where  $x$  = low socioeconomic proficient and above state in the West census-defined region and  $y$  = low socioeconomic proficient and above NAEP in the West census-defined region, indicated that the linear model was statistically significant in predicting the low socioeconomic NAEP proficient and above from the low socioeconomic state proficient and above data in the West census-defined region.

7. A regression analysis using the linear model:

$\hat{y} = .583 + .157x$ , where  $x$  = nonwhite proficient and above state in the Midwest census-defined region and  $y$  = nonwhite proficient and above NAEP in the Midwest census-defined region, indicated that the linear model was statistically significant in predicting the nonwhite state proficient and above from the nonwhite NAEP proficient and above data in the Midwest census-defined region.

8. A regression analysis using the linear model:

$\hat{y} = -1.955 + .501x$ , where  $x$  = nonwhite proficient and above state in the Northeast census-defined region and  $y$  = nonwhite proficient and above NAEP in the Northeast census-defined region, indicated that the linear model was statistically significant in predicting the nonwhite NAEP proficient and above from the nonwhite state proficient and above data in the Northeast census-defined region.

9. A regression analysis using the linear model:

$\hat{y} = 1.861 + .150x$ , where  $x$  = nonwhite proficient and above state in the South census-defined region and  $y$  = nonwhite proficient and above NAEP in the South census-refined region, indicated that the linear model was statistically significant in predicting the nonwhite NAEP proficient and above from the nonwhite state proficient and above data in the South census-defined region.

10. A regression analysis using the linear model:

$\hat{y} = .733 + .272x$ , where  $x$  = nonwhite proficient and above state in the West census-defined region and  $y$  = nonwhite proficient and above NAEP in the West census-defined region, indicated that the linear model was statistically significant in predicting the nonwhite NAEP proficient and above from the nonwhite state proficient and above data in the West census-defined region.

### Implications for Practice

The review of literature demonstrated that the federal government increased accountability and assessment measures as a means to influence states' educational policies. Additionally, increased transparency of student performance data forced states to display their data for all stakeholders to view. The lack of uniform accountability systems, between states, has created difficulty for educators, policy-makers and the general public to make state-by-state comparisons in the area of reading. Through this study, the following recommendations can be made but are not limited to:

1. Investigate using a national assessment system which could promote a better way for comparisons to be made on student performance.
2. Create and implement national standards that can be measured by the NAEP. By creating national standards and using the NAEP to assess them, performance of mobile students could be gauged on a common basis.

3. Provide professional development for local level administrators and teachers to promote understanding of how to analyze state scores in relationship to NAEP scores in order to understand national comparisons.
4. Institute the NAEP as a means to reduce expenses associated with state standardized testing.

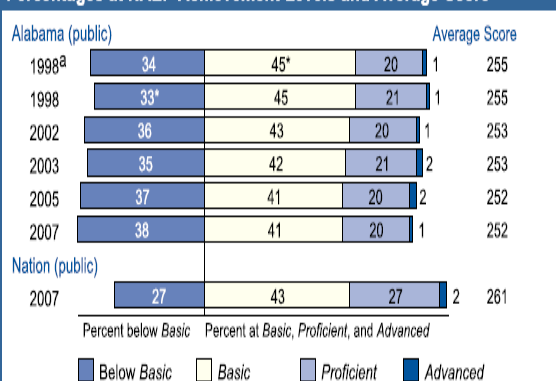
### Future Research

The following recommendations for future research should be considered but not limited to:

1. More research should be conducted to determine if national standards and a national assessment system is cost effective to implement.
2. More research should be conducted determine the relationship that exists, if any, among state standards by regions.
3. Further research needs to be conducted to determine why the low socioeconomic West census-defined region was able to predict NAEP Proficient and above scores from state proficient and above scores.
4. Further research needs to be conducted to determine why the Low socioeconomic Census-Defined Midwest, South and North regions were not able to predict NAEP proficient and above scores from state proficient and above scores.

APPENDIX A: NCES 2007 STATE SNAPSHOT REPORT (ALABAMA)

The National Assessment of Educational Progress (NAEP) assesses reading in three content areas in grade 8: reading for literary experience, to gain information, and to perform a task. The NAEP reading scale ranges from 0 to 500.

Overall Reading Results for Alabama		Percentages at NAEP Achievement Levels and Average Score																																																							
<ul style="list-style-type: none"> <li>In 2007, the average scale score for eighth-grade students in Alabama was 252. This was not significantly different from their average score in 2005 (252) and was not significantly different from their average score in 1998 (255).<sup>1</sup></li> <li>Alabama's average score (252) in 2007 was lower than that of the nation's public schools (261).</li> <li>Of the 52 states and other jurisdictions that participated in the 2007 eighth-grade assessment, students' average scale score in Alabama was higher than that in 1 jurisdiction, not significantly different from those in 7 jurisdictions, and lower than those in 43 jurisdictions.<sup>2</sup></li> <li>The percentage of students in Alabama who performed at or above the NAEP <i>Proficient</i> level was 21 percent in 2007. This percentage was not significantly different from that in 2005 (22 percent) and was not significantly different from that in 1998 (22 percent).</li> <li>The percentage of students in Alabama who performed at or above the NAEP <i>Basic</i> level was 62 percent in 2007. This percentage was not significantly different from that in 2005 (63 percent) and was smaller than that in 1998 (67 percent).</li> </ul>		 <p>Alabama (public) Average Score</p> <table border="1"> <thead> <tr> <th>Year</th> <th>Below Basic</th> <th>Basic</th> <th>Proficient</th> <th>Advanced</th> <th>Average Score</th> </tr> </thead> <tbody> <tr> <td>1998<sup>a</sup></td> <td>34</td> <td>45*</td> <td>20</td> <td>1</td> <td>255</td> </tr> <tr> <td>1998</td> <td>33*</td> <td>45</td> <td>21</td> <td>1</td> <td>255</td> </tr> <tr> <td>2002</td> <td>36</td> <td>43</td> <td>20</td> <td>1</td> <td>253</td> </tr> <tr> <td>2003</td> <td>35</td> <td>42</td> <td>21</td> <td>2</td> <td>253</td> </tr> <tr> <td>2005</td> <td>37</td> <td>41</td> <td>20</td> <td>2</td> <td>252</td> </tr> <tr> <td>2007</td> <td>38</td> <td>41</td> <td>20</td> <td>1</td> <td>252</td> </tr> <tr> <td colspan="6"><b>Nation (public)</b></td> </tr> <tr> <td>2007</td> <td>27</td> <td>43</td> <td>27</td> <td>2</td> <td>261</td> </tr> </tbody> </table> <p> <small>Percent below Basic    Percent at Basic, Proficient, and Advanced</small>  <small>■ Below Basic    □ Basic    ■ Proficient    ■ Advanced</small> </p> <p><sup>a</sup> Accommodations were not permitted for this assessment.</p> <p>NOTE: The NAEP grade 8 reading achievement levels correspond to the following scale points: <i>Below Basic</i>, 242 or lower; <i>Basic</i>, 243–280; <i>Proficient</i>, 281–322; <i>Advanced</i>, 323 or above.</p>		Year	Below Basic	Basic	Proficient	Advanced	Average Score	1998 <sup>a</sup>	34	45*	20	1	255	1998	33*	45	21	1	255	2002	36	43	20	1	253	2003	35	42	21	2	253	2005	37	41	20	2	252	2007	38	41	20	1	252	<b>Nation (public)</b>						2007	27	43	27	2	261
Year	Below Basic	Basic	Proficient	Advanced	Average Score																																																				
1998 <sup>a</sup>	34	45*	20	1	255																																																				
1998	33*	45	21	1	255																																																				
2002	36	43	20	1	253																																																				
2003	35	42	21	2	253																																																				
2005	37	41	20	2	252																																																				
2007	38	41	20	1	252																																																				
<b>Nation (public)</b>																																																									
2007	27	43	27	2	261																																																				

Reporting groups	Percent of students	Average score	Percent below Basic	Percent of students at or above		Percent Advanced
				Basic	Proficient	
Male	50	247	43	57	18	1
Female	50	257	33	67	25	2
White	60	261	27	73	29	2
Black	36	236	57	43	9	#
Hispanic	3	250	39	61	20	1
Asian/Pacific Islander	1	‡	‡	‡	‡	‡
American Indian/Alaska Native	#	‡	‡	‡	‡	‡
Eligible for National School Lunch Program	49	241	50	50	11	#
Not eligible for National School Lunch Program	51	263	25	75	31	2

Source: Institute of Educational Sciences (IES), National Center for Education Statistics (NCES) 2007-497AL8, State Snapshot Report



APPENDIX B: SY 2006-2007 CSPR REPORT (ALABAMA)

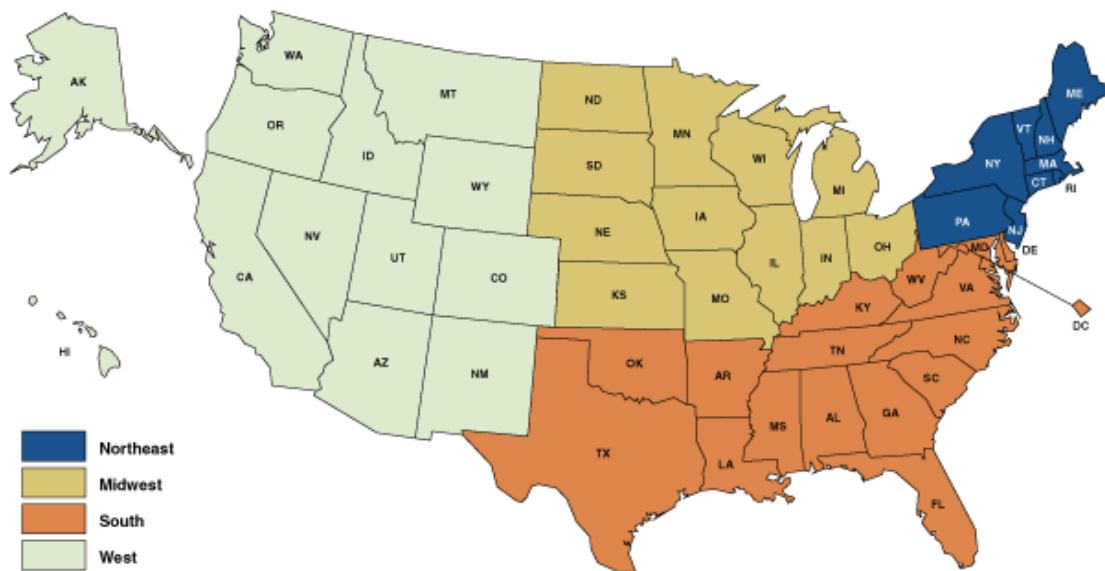
1.3.12 Student Academic Achievement in Reading/Language Arts - Grade 8

Grade 8	# Students Who Completed the Assessment and for Whom a Proficiency Level Was Assigned	# Students Scoring at or Above Proficient	Percentage of Students Scoring at or Above Proficient
All students	58137	41742	71.8
American Indian or Alaska Native	552	447	81.0
Asian or Pacific Islander	544	447	82.2
Black, non-Hispanic	21312	12535	58.8
Hispanic	1549	875	56.5
White, non-Hispanic	34102	27388	80.3
Children with disabilities ( <i>IDEA</i> )	7107	1802	25.4
Limited English proficient (LEP) students	996	362	36.3
Economically disadvantaged students	30283	18337	60.6
Migratory students	148	82	55.4
Male	29798	19594	65.8
Female	28339	22148	78.2
Comments: Increase in Hispanic and Asian population due to new industry and influx of hispanics into service industries			

Source – Initially prepopulated by ED*Facts* in file N/X076, N/X077, or N/X078 that are data group 584, category sets A, B, C, D, E, and F. If the SEA has additional racial/ethnic groups or combinations of racial/ethnic groups in its accountability plan under NCLB, the SEA will report the above data for those groups through the online C*SPR* collection tool.

Note: The addition of the total number of students with an assigned proficiency level is a new data collection for the SY 2006-07 C*SPR*.

APPENDIX C: THE FOUR CENSUS-DEFINED REGIONS OF NAEP



- *Northeast:* Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont;
- *South:* Alabama, Arkansas, Delaware, District of Columbia, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia, and West Virginia;
- *Midwest:* Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, and Wisconsin; and
- *West:* Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington, and Wyoming.

Source: U.S. Department of Education, Institute of Education Sciences, National Center For Education Statistics, National Assessment of Educational Progress (NAEP), Census-Defined Regions of NAEP

APPENDIX D: SPSS WORKSHEET

State	Census-Defined Region	State Proficient and Above (%)	NAEP Proficient and Above (%)	Low Socio-economic State (%)	Low Socio-economic NAEP (%)	Nonwhite State (%)	Nonwhite NAEP (%)
AL	South	72	21	61	11	25	3
AK	West	79	27	68	14	27	14
AZ	West	63	24	47	11	27	7
AR	South	63	25	52	15	14	3
CA	West	42	22	26	11	23	10
CO	West	87	34	75	18	28	7
CT	Northeast	75	38	51	14	17	5
DE	South	78	30	66	18	32	8
FL	South	49	28	35	17	18	10
GA	South	89	26	83	14	46	7
HI	West	60	20	48	13	50	13
ID	West	86	32	78	22	12	2
IL	Midwest	81	29	71	15	31	6
IN	Midwest	68	31	52	16	9	2
IA	Midwest	73	35	56	22	7	2
KS	Midwest	81	35	68	20	16	3
KY	South	64	28	53	17	6	2
LA	South	59	19	49	12	22	4
ME	Northeast	65	37	48	26	3	0
MD	South	69	33	50	17	30	10
MA	Northeast	75	43	54	20	14	5
MI	Midwest	72	28	57	12	15	2
MN	Midwest	64	37	43	21	9	3
MS	South	52	17	39	10	20	4
MO	Midwest	43	32	27	18	5	2
MT	West	79	39	66	24	9	2
NE	Midwest	91	35	84	21	17	3
NV	West	57	22	43	12	25	7
NH	Northeast	66	37	42	25	3	0
NJ	Northeast	72	39	51	16	25	12
NM	West	56	18	47	10	34	8
NY	Northeast	57	33	40	19	20	8
NC	South	88	28	81	14	33	5
ND	Midwest	76	32	64	20	7	1
OH	Midwest	80	36	66	16	13	2
OK	South	78	26	69	18	28	7
OR	West	68	34	54	21	13	5
PA	Northeast	74	36	56	20	13	5
RI	Northeast	58	27	40	12	11	2
SC	South	35	25	20	11	9	4
SD	Midwest	78	37	65	25	9	2
TN	South	92	26	88	14	27	3
TX	South	88	28	82	15	52	10
UT	West	81	30	69	21	12	3
VT	Northeast	65	42	47	25	3	0
VA	South	80	34	65	16	27	8
WA	West	67	34	52	20	17	7
WV	South	80	23	72	15	5	1
WI	Midwest	84	34	70	16	15	3
WY	West	71	33	57	22	8	2

Source: United States Department of Education's SY 2006-2007 Consolidated State Performance Report. U.S. Department of Education, Institute of Education Sciences, National Center For Education Statistics, National Assessment of Educational Progress (NAEP), various years, 1998-2007 Reading Assessments

APPENDIX E: STATE AND NAEP DATA (PROFICIENT AND ABOVE)

State	State Proficient and Above (%)	NAEP Proficient and Above (%)
AL	72	21
AK	79	27
AZ	63	24
AR	63	25
CA	42	22
CO	87	34
CT	75	38
DE	78	30
FL	49	28
GA	89	26
HI	60	20
ID	86	32
IL	81	29
IN	68	31
IA	73	35
KS	81	35
KY	64	28
LA	59	19
ME	65	37
MD	69	33
MA	75	43
MI	72	28
MN	64	37
MS	52	17
MO	43	32
MT	79	39
NE	91	35
NV	57	22
NH	66	37
NJ	72	39
NM	56	18
NY	57	33
NC	88	28
ND	76	32
OH	80	36
OK	78	26
OR	68	34
PA	74	36
RI	58	27
SC	35	25
SD	78	37
TN	92	26
TX	88	28
UT	81	30
VT	65	42
VA	80	34
WA	67	34
WV	80	23
WI	84	34
WY	71	33

Source: United States Department of Education's SY 2006-2007 Consolidated State Performance Report. U.S. Department of Education, Institute of Education Sciences, National Center For Education Statistics, National Assessment of Educational Progress (NAEP), various years, 1998-2007 Reading Assessments



APPENDIX F: STATE AND NAEP BY CENSUS-DEFINED REGION

State	Census-Defined Region	State Proficient and Above (%)	NAEP Proficient and Above (%)
AL	South	72	21
AK	West	79	27
AZ	West	63	24
AR	South	63	25
CA	West	42	22
CO	West	87	34
CT	Northeast	75	38
DE	South	78	30
FL	South	49	28
GA	South	89	26
HI	West	60	20
ID	West	86	32
IL	Midwest	81	29
IN	Midwest	68	31
IA	Midwest	73	35
KS	Midwest	81	35
KY	South	64	28
LA	South	59	19
ME	Northeast	65	37
MD	South	69	33
MA	Northeast	75	43
MI	Midwest	72	28
MN	Midwest	64	37
MS	South	52	17
MO	Midwest	43	32
MT	West	79	39
NE	Midwest	91	35
NV	West	57	22
NH	Northeast	66	37
NJ	Northeast	72	39
NM	West	56	18
NY	Northeast	57	33
NC	South	88	28
ND	Midwest	76	32
OH	Midwest	80	36
OK	South	78	26
OR	West	68	34
PA	Northeast	74	36
RI	Northeast	58	27
SC	South	35	25
SD	Midwest	78	37
TN	South	92	26
TX	South	88	28
UT	West	81	30
VT	Northeast	65	42
VA	South	80	34
WA	West	67	34
WV	South	80	23
WI	Midwest	84	34
WY	West	71	33

Source: United States Department of Education's SY 2006-2007 Consolidated State Performance Report. U.S. Department of Education, Institute of Education Sciences, National Center For Education Statistics, National Assessment of Educational Progress (NAEP), various years, 1998-2007 Reading Assessment

APPENDIX G: LOW-SOCIOECONOMIC CENSUS-DEFINED REGION

State	Census-Defined Region	Low Socioeconomic State (%)	Low Socioeconomic NAEP (%)
AL	South	61	11
AK	West	68	14
AZ	West	47	11
AR	South	52	15
CA	West	26	11
CO	West	75	18
CT	Northeast	51	14
DE	South	66	18
FL	South	35	17
GA	South	83	14
HI	West	48	13
ID	West	78	22
IL	Midwest	71	15
IN	Midwest	52	16
IA	Midwest	56	22
KS	Midwest	68	20
KY	South	53	17
LA	South	49	12
ME	Northeast	48	26
MD	South	50	17
MA	Northeast	54	20
MI	Midwest	57	12
MN	Midwest	43	21
MS	South	39	10
MO	Midwest	27	18
MT	West	66	24
NE	Midwest	84	21
NV	West	43	12
NH	Northeast	42	25
NJ	Northeast	51	16
NM	West	47	10
NY	Northeast	40	19
NC	South	81	14
ND	Midwest	64	20
OH	Midwest	66	16
OK	South	69	18
OR	West	54	21
PA	Northeast	56	20
RI	Northeast	40	12
SC	South	20	11
SD	Midwest	65	25
TN	South	88	14
TX	South	82	15
UT	West	69	21
VT	Northeast	47	25
VA	South	65	16
WA	West	52	20
WV	South	72	15
WI	Midwest	70	16
WY	West	57	22

Source: United States Department of Education's SY 2006-2007 Consolidated State Performance Report. U.S. Department of Education, Institute of Education Sciences, National Center For Education Statistics, National Assessment of Educational Progress (NAEP), various years, 1998-2007 Reading Assessments

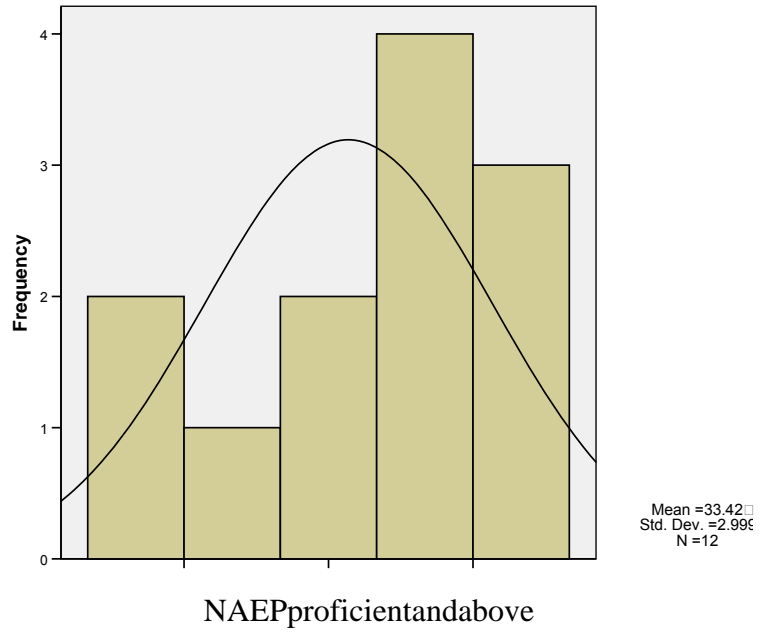
APPENDIX H: NONWHITE CENSUS-DEFINED REGION

State	Census-Defined Region	Nonwhite State (%)	Nonwhite NAEP (%)
AL	South	25	3
AK	West	27	14
AZ	West	27	7
AR	South	14	3
CA	West	23	10
CO	West	28	7
CT	Northeast	17	5
DE	South	32	8
FL	South	18	10
GA	South	46	7
HI	West	50	13
ID	West	12	2
IL	Midwest	31	6
IN	Midwest	9	2
IA	Midwest	7	2
KS	Midwest	16	3
KY	South	6	2
LA	South	22	4
ME	Northeast	3	0
MD	South	30	10
MA	Northeast	14	5
MI	Midwest	15	2
MN	Midwest	9	3
MS	South	20	4
MO	Midwest	5	2
MT	West	9	2
NE	Midwest	17	3
NV	West	25	7
NH	Northeast	3	0
NJ	Northeast	25	12
NM	West	34	8
NY	Northeast	20	8
NC	South	33	5
ND	Midwest	7	1
OH	Midwest	13	2
OK	South	28	7
OR	West	13	5
PA	Northeast	13	5
RI	Northeast	11	2
SC	South	9	4
SD	Midwest	9	2
TN	South	27	3
TX	South	52	10
UT	West	12	3
VT	Northeast	3	0
VA	South	27	8
WA	West	17	7
WV	South	5	1
WI	Midwest	15	3
WY	West	8	2

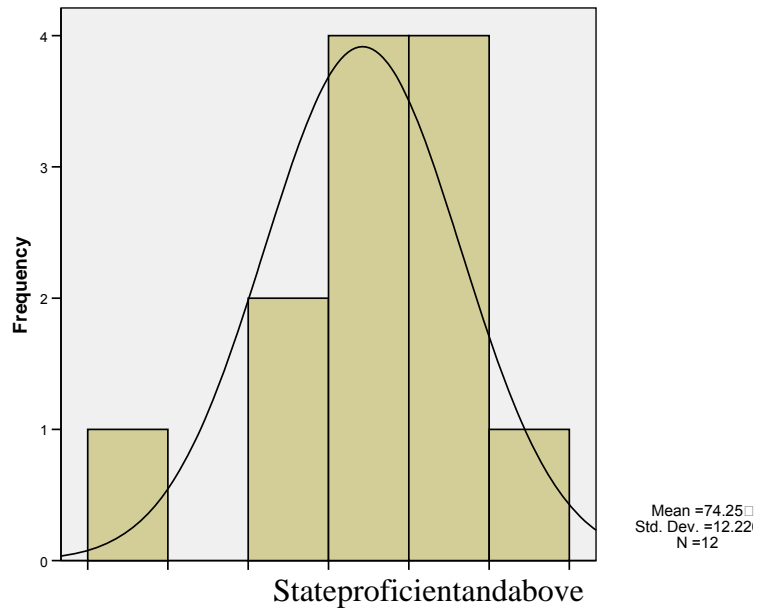
Source: United States Department of Education's SY 2006-2007 Consolidated State Performance Report. U.S. Department of Education, Institute of Education Sciences, National Center For Education Statistics, National Assessment of Educational Progress (NAEP), various years, 1998-2007 Reading Assessments

APPENDIX I: HISTOGRAMS WITH NORMAL CURVES (MIDWEST REGION)

### NAEP Proficient and Above



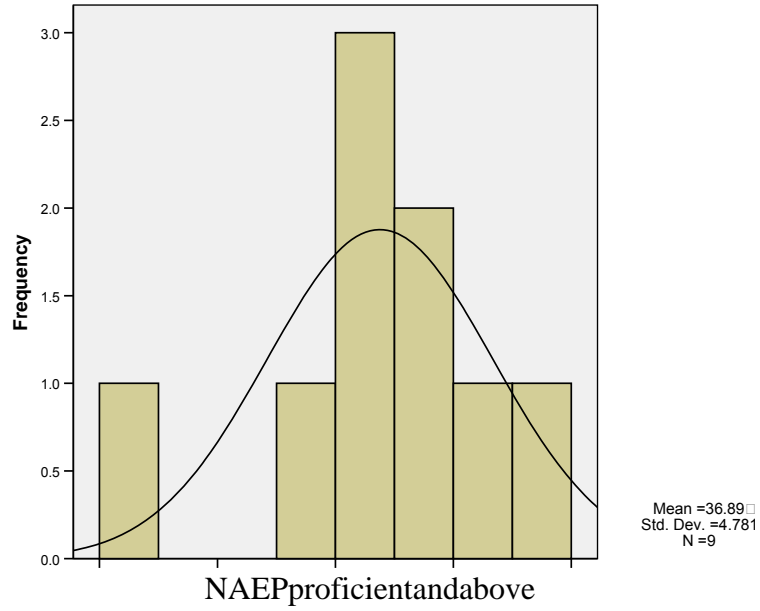
### State Proficient and Above



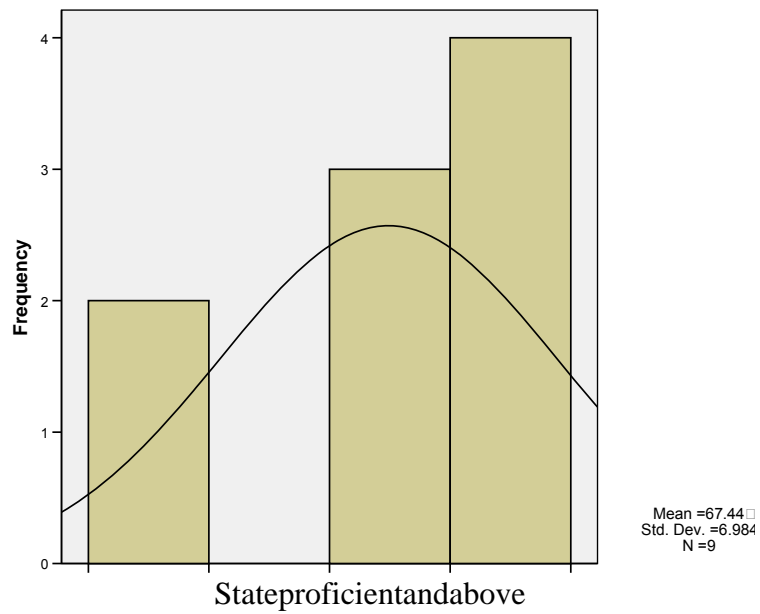


APPENDIX J: HISTOGRAMS WITH NORMAL CURVES (N.E. REGION)

### NAEP Proficient and Above

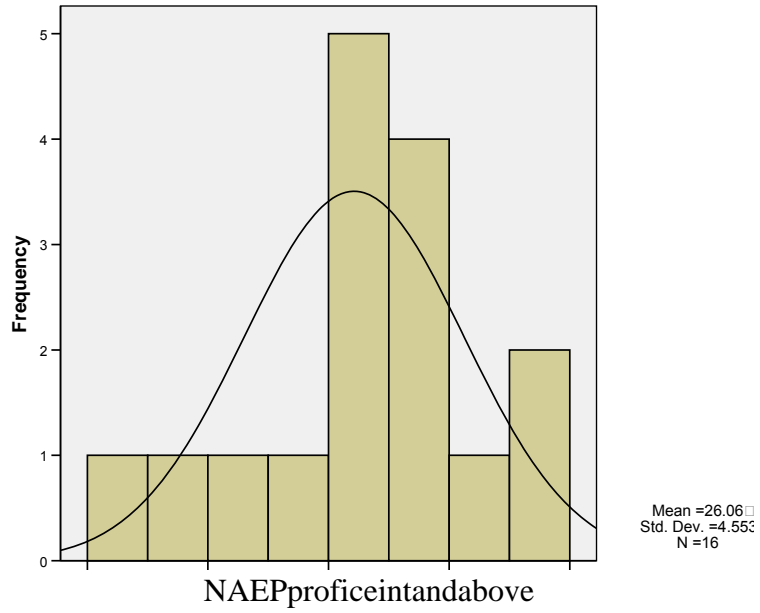


### State Proficient and Above

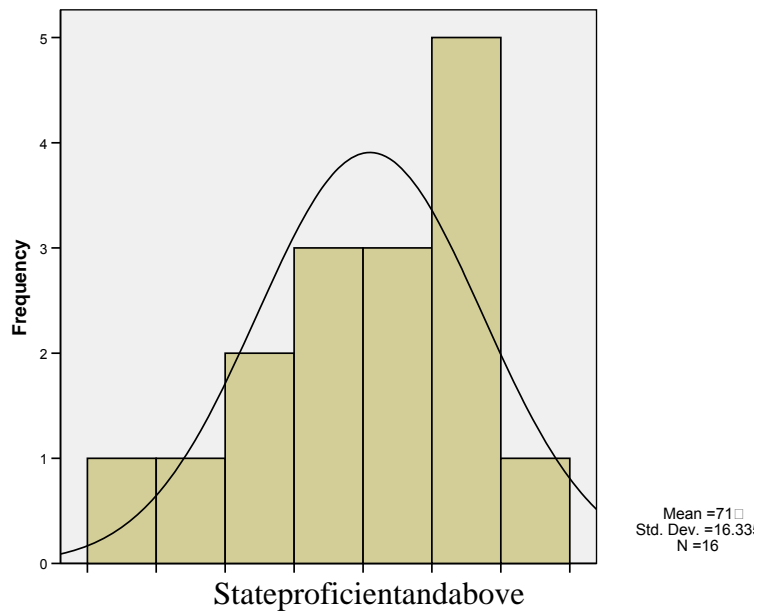


### APPENDIX K: HISTOGRAMS WITH NORMAL CURVES (SOUTH REGION)

### NAEP Proficient and Above

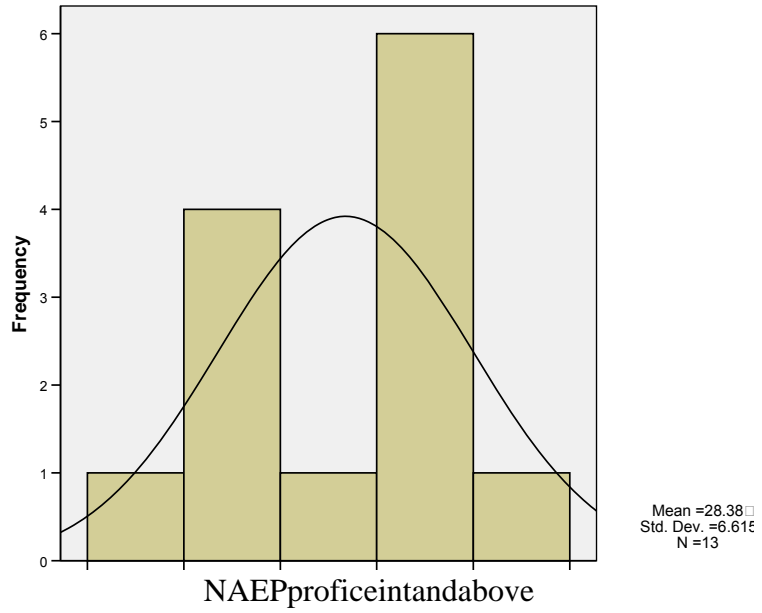


### State Proficient and Above

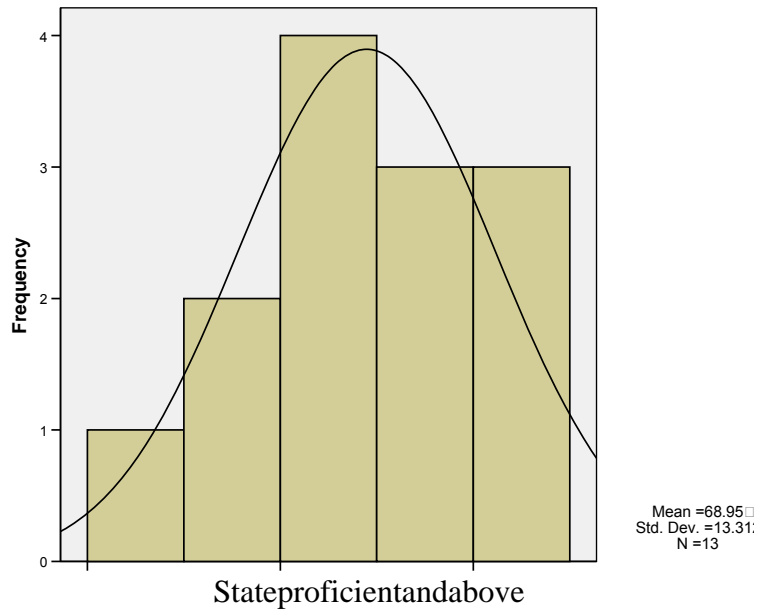


APPENDIX L: HISTOGRAMS WITH NORMAL CURVES (WEST REGION)

### NAEP Proficient and Above



### State Proficient and Above



APPENDIX M: STUDENT PERFORMANCE ON STATE AND NAEP EXAMS

### Average Student Scores

Census-Defined Region	Regional Averages		Low Socioeconomic Averages (%)		Nonwhite Averages (%)	
	State (%)	NAEP (%)	State (%)	NAEP (%)	State (%)	NAEP (%)
Midwest	74	33	60	19	11	2
Northeast	67	37	48	20	12	4
South	71	26	60	15	25	6
West	69	28	56	17	22	7

Source: United States Department of Education's SY 2006-2007 Consolidated State Performance Report. U.S. Department of Education, Institute of Education Sciences, National Center For Education Statistics, National Assessment of Educational Progress (NAEP), various years, 1998-2007 Reading Assessment

### Average Student Performance of All Regions

	Regional Averages		Low Socioeconomic Averages		Nonwhite Averages	
	State (%)	NAEP (%)	State (%)	NAEP (%)	State (%)	NAEP (%)
Total	71	34	57	17	19	5

Source: United States Department of Education's SY 2006-2007 Consolidated State Performance Report. U.S. Department of Education, Institute of Education Sciences, National Center For Education Statistics, National Assessment of Educational Progress (NAEP), various years, 1998-2007 Reading Assessment

APPENDIX N: UCF SUBMISSION OUTCOME LETTER





University of Central Florida Institutional Review Board  
Office of Research & Commercialization  
12201 Research Parkway, Suite 501  
Orlando, Florida 32826-3246  
Telephone: 407-823-2901, 407-882-2901 or 407-882-2276  
[www.research.ucf.edu/compliance/irb.html](http://www.research.ucf.edu/compliance/irb.html)

### Not Human Subjects Research

From : UCF Institutional Review Board  
FWA00000351, Exp. 6/24/11, IRB00001138

To : William R. Gordon

Date : October 23, 2008

IRB Number: SBE-08-05860

Study Title: A COMPARISON OF EIGHTH GRADE READING SCORES BY STATE AND BY THE FOUR REGIONS IDENTIFIED BY NAEP

Dear Researcher:

After reviewing the materials that you have submitted, and the fact that the data you are using is publicly available, the UCF Institutional Review Board has determined that your study does not fit the definition of human subjects research.

Therefore, IRB review is not needed.

Thank you for your time in resolving this issue. Please continue to submit applications that involve human subject activities that could potentially involve human subjects as research participants.

On behalf of Tracy Dietz, Ph.D., UCF IRB Chair, this letter is signed by:

Signature applied by Joanne Muratori on 10/23/2008 01:22:03 PM EDT

A handwritten signature in cursive script that reads 'Joanne Muratori'.

IRB Coordinator

## LIST OF REFERENCES

- 107<sup>th</sup> Congress. (2002, January 8). Public Law – 107 -110. Retrieved August 18, 2008, from <http://www.ed.gov/policy/elsec/leg/esea02/107-110.pdf>
- Ad Hoc Committee on Confirming Test Results. (2002). Using the national assessment of educational progress to confirm state test results. Washington, DC: National Governing Assessment Governing Board. Retrieved July, 15, 2008, from <http://nagb.org>
- Beaton, A. E., & Zwick, R. (1992). Overview of the national assessment of educational progress. *Journal of Educational Statistics*, 17(2, Special Issue: National Assessment of Educational Progress), 95-109.
- Blum, R. (2000). Standards-based reform: can it make a difference for students? *Peabody Journal of Education*, 75(4), 90–113.
- Bowers, J. (1991, Winter). Evaluating testing programs at state and local levels. *Theory into Practice*, 30(1), 52-60.
- Bracey, G. (2007, October). The first time “everything changed”: The 17<sup>th</sup> bracey report on the condition of education. *Phi Delta Kappan*, 89(2), 119-136.
- Brown, F. (2004, Summer). The first serious implementation of brown: The 1964 civil rights act and beyond. *The Journal of Negro Education*, 73(3), 182-190.
- Electronic Summary of *A Nation At Risk*. (2001). Retrieved June 13, 2008, from <http://www.ed.gov/pubs/NatAtRisk/intro.html>
- Electronic Summary of The Elementary and Secondary Act of 1965. (2008). Retrieved June 9, 2008, from [http://wps.prenhall.com/wps/media/objects/2355/2412111/Documents\\_Library/ea1965.html](http://wps.prenhall.com/wps/media/objects/2355/2412111/Documents_Library/ea1965.html)
- Goertz, M. ( 2001). The federal role in defining “adequate yearly progress”: The flexibility/accountability tradeoff. Washington, DC: National Institute on Educational Governance, Finance, Policymaking, and Management, 1-16.
- Hoff, D. (1998, January). Controversial goals 2000 to face new uncertainties. *Education Week*, 17(19).

- Institute of Education Sciences. (2007, October). Final report on the national assessment of title I: Summary of key findings. Washington, DC: U.S. Department of Education. Retrieved September 7, 2008, from [http://ies.ed.gov/ncee/pdf/20084014\\_rev.pdf](http://ies.ed.gov/ncee/pdf/20084014_rev.pdf)
- Jones, L., & Olkin, I. (2004). *The nation's report card: Evolution and perspectives*. Bloomington, IN: Phi Delta Kappa Educational Foundation.
- Kantor, H. (1991, November). Education, social reform, and the state: ESEA and federal education policy in the 1960s. *American Journal of Education, 100*(1), 47–83.
- Katz, M. B. (1968). *The irony of early school reform*. Cambridge, MA: Harvard Press.
- Lee, J. (2007). Do national and state assessments converge for educational accountability? A meta-analytic synthesis of multiple measures in maine and kentucky. *Applied Measurement in Education, 20*(2), 171-203.
- Linn, R. (2000, March). Assessments and accountability. *Educational Researcher, 29*(2), 4-16.
- Linn, R. (2003, October). 2003 presidential address. Accountability: responsibility and reasonable expectations. *Educational Researcher, 32*(7), 3-13.
- Linn, R, Baker, E., & Bettebenner, D. (2002, August). Accountability systems: Implications of requirements of the No Child Left Behind Act of 2001. *Educational Researcher, 31*(6), 3-16.
- Lunenburg, F. C., & Ornstein, A.C. (2004). *Educational administration: concepts and practices*. Belmont, CA: Thomson Learning.
- Madaus, G. F. (1985). Public policy and the testing profession: you've never had it so good? *Educational Measurement: Issues and Practice, 4*(4), 5-11.
- Mondale, S., & Patton, S. (2001). *The story of American public education*. Boston: Beacon Press.
- National Assessment Governing Board. (2008). Reading framework for the 2005 National Assessment of Educational Progress. Retrieved September 15, 2008, from <http://nagb.org>
- National Center for Education Statistics. (2007). Mapping 2005 state proficiency standards on the NAEP scales. NCES 2007-482, U.S. Department of Education, Washington DC: Author. Retrieved September 11, 2008, from <http://nces.ed.gov>

- Ohanian, S. (2000, January). Goals 2000: What's in a name? *Phi Delta Kappan*, 81(5), 344-355.
- Reese, W. (2007, Summer). Why Americans love to reform the public schools. *Education Horizons*, 85(4), 217-231.
- Reisner, E., et al. (1982, April). *Assessment of the title I evaluation and reporting system*. United States Department of Education, Washington, DC: Office of Planning, Budget and Evaluation, 1-141.
- Resnick, D. P. (1980). Minimum competency testing historically considered. *Review of Research in Education*, 8, 3-29.
- Rickover, H. G. (1959). *Education and Freedom*. New York: Dutton & Company.
- Schafer, W., Liu, M., & Wang, H. (2007, August). Content and grade trends in state assessments and NAEP. *Practical Assessment, Research and Evaluation*, 12(9), 1-25.
- Shaul, M., & Ganson, H. (2005). The no child left behind act of 2001: The federal government's role in strengthening accountability for student performance. *Review of Research in Education*, 29, 151-165.
- Stoneberg, B. D. (2007, March). *The valid use of NAEP achievement level scores confirm state testing results in No Child Left Behind Act*. Paper presented at the National Assessment of Educational Progress (NAEP) State Service Center Spring Assessment Workshop, Bethesda, MD.
- Swanson, C. B., & Stevenson, D. L. (2002, Spring). Standards-based reform in practice: evidence on state policy and classroom instruction from the NAEP state assessments. *Educational Evaluation and Policy Analysis*, 24(1), 1-27.
- Tallmadge, G. K., & Wood, C. T. (1981). *Users guide: ESEA title I evaluation and reporting system*. Mountain View, CA: RMC Corporation.
- Thompson, M. (2004, February). *No child left behind act of 2001: Implementation of the reauthorized ESEA*. PowerPoint presented at the meeting of the National At-Risk Education Network, Panama City, Florida.
- Tyack, D. (2000, November). *Reflections on the history of U.S. education*. *Educational Researcher*, 29(8), 19-20.
- Tyack, D., & Cuban, L. (1995). *Tinkering toward utopia: A century of public school reform*. Cambridge, Mass.: Harvard University Press.

- United States Department of Education. (1994a). *Improving America's Schools Act of 1994*. Retrieved October 15, 2008, from <http://www.ed.gov/legislation/ESEA/index.html>
- United States Department of Education. (1994b). *Goals 2000: Educate America Act*. Retrieved October 15, 2008, from <http://www.ed.gov/legislation/GOALS2000/TheACT/intro.html>
- United States Department of Education. (1998, April). *Executive summary: Reforming education to improve student achievement*. Retrieved October 13, 2008, from <http://www.ed.gov/pubs/G2KReforming/g2exec.html>
- United States Department of Education. (1999, April). *Federal education legislation enacted in 1994. An evaluation of implementation and impact*. Washington, DC: Planning and Evaluation Services.
- United States Department of Education. (2004a). *Four pillars of NCLB*. Retrieved September 7, 2008, from <http://www.ed.gov/nclb/overview/intro/4pillars.html>
- United States Department of Education. (2004b). *Title I- Improving the academic achievement of the disadvantaged*. Section 1002 (a)(6). Retrieved September 6, 2008, from <http://www.ed.gov/policy/elsec/leg/esea02/pg1.html>
- United States Department of Education. (2004c). *Title VI - Flexibility and accountability*. Retrieved October 2, 2008, from <http://www.ed.gov/policy/elsec/leg/esea02/pg87.html>
- United States Department of Education. (2004d). *Executive summary of the no child left behind act of 2001*. Retrieved June 8, 2008, from <http://www.ed.gov./print/nclb/overview/intro/execsumm.html>
- United States Department of Education. ( 2005). *Title I-Improving the academic achievement of the disadvantaged*. Sections 1111 (b)(2)(A); 1111 (b)(2)(C)(1). Retrieved September 6, 2008, from <http://www.ed.gov/policy/elsec/leg/esea02/pg2.html#sec1111>
- United States Department of Education. (2007a). *National high school transcript study: Census-defined regions*. Retrieved October 18, 2008, from <http://nces.ed.gov/nationsreportcard/hsts/tabulations/regions.asp>
- United States Department of Education. (2007b). *No child left behind: A desktop reference*. Retrieved October 5, 2008, from [http://www.ed.gov/admins/lead/account/nclbreference/page\\_pg54.html](http://www.ed.gov/admins/lead/account/nclbreference/page_pg54.html)

- United States Department of Education. (2008a). *The nation's report card*. Retrieved October 15, 2008, from <http://nces.ed.gov/nationsreportcard/faq.asp#ques13>
- United States Department of Education. (2008b). *An introduction to national assessment of educational progress*. Jessup, MD. Retrieved September 13, 2008, from [http://nces.ed.gov/nationsreportcard/pdf/about/introduction\\_to\\_naep\\_2008.pdf](http://nces.ed.gov/nationsreportcard/pdf/about/introduction_to_naep_2008.pdf)
- Waltman, K. (1997, Summer). Using performance standards to link statewide achievement results to NAEP. *Journal of Education Measurement*, 34 (2) 101-121.
- Wong, K., & Meyer, S. (1998). An overview of title I school-wide programs: federal legislative expectations. Outlined in the improving America's schools act of 1994 amendments to title I of the elementary on secondary act of 1965. Washington, DC: Office of Educational Research and Improvement